

Educación para todos no es un proyecto lucrativo, sino un esfuerzo colectivo de estudiantes y profesores de la UNAM para facilitar el acceso a los materiales necesarios para la educación de la mayor cantidad de gente posible. Pensamos editar en formato digital libros que por su alto costo, o bien porque ya no se consiguen en bibliotecas y librerías, no son accesibles para todos.

Invitamos a todos los interesados en participar en este proyecto a sugerir títulos, a prestarnos los textos para su digitalización y a ayudarnos en toda la labor técnica que implica su reproducción. El nuestro, es un proyecto colectivo abierto a la participación de cualquier persona y todas las colaboraciones son bienvenidas.

Nos encuentras en los Talleres Estudiantiles de la Facultad de Ciencias y puedes ponerte en contacto con nosotros a la siguiente dirección de correo electrónico:

eduktodos@gmail.com http://eduktodos.dyndns.org

Calculus

Tom M. Apostol

CALCULUS

VOLUMEN II

Cálculo con funciones de varias variables y álgebra lineal, con aplicaciones a las ecuaciones diferenciales y a las probabilidades

Segunda edición



EDITORIAL REVERTÉ, S. A.
Barcelona-Bogotá-Buenos Aires-Caracas-México

Título de la obra original:
CALCULUS, Multi-Variable Calculus and Linear Algebra, With Applications to Differential Equations and Probability

Edición original en lengua inglesa publicada por:
Blaisdell Publishing Company, Waltham, Massachusetts

Versión española por:
Dr. D. Francisco Vélez Cantarell
Profesor de la Universitat de Barcelona

Revisada por:
Dr. D. Enrique Linés Escardó
Catedrático de la Facultad de Ciencias de la Universidad de Madrid

Propiedad de:
EDITORIAL REVERTÉ, S.A. y
Loreto, 13-15, Local B
08029 Barcelona
Tel: (34) 93 419 33 36
Fax: (34) 93 419 51 89
E-mail: reverte@reverte.com
Internet: http://www.reverte.com

REVERTÉ EDICIONES, S.A. DE CV
Río Pánuco 141 Col. Cuauhtémoc
C.P. 06500 México, D.F.
Tel: 55-33-56-58 al 60
Fax: 55-14-67-99
E-mail: resavbp@data.net.mx

2ª. EDICIÓN

Edición en español

7ª REIMPRESIÓN: MARZO DE 2002

ISBN: 84-291-5001-3 (Obra completa) España
ISBN: 84-291-5003-X (Tomo 2)

ISBN: 698-6708-12-X (Obra completa) México
ISBN: 698-6708-11-1 (Tomo 2)
Depósito legal: B-13143-2002

Impreso por Domingraf Impressors
Pol. Ind. Can Magarola
Pje. Autopista, Nave 12
08100 Mollet del Vallés (Barcelona)

a

Jane y Stephen

PRÓLOGO

Este libro es una continuación de mi Calculus, *volumen I, segunda edición. El presente volumen fue escrito con el mismo plan fundamental que inspiró al primero. Un adecuado enfoque hacia la técnica se combina con un riguroso desarrollo teórico. Se ha procurado hacer llegar al estudiante el espíritu de la matemática moderna sin exagerar el formalismo. Como en el volumen I, se han incluido comentarios de tipo histórico para hacer vivir al lector la evolución de las ideas.*

El segundo volumen está dividido en tres partes, tituladas: Análisis lineal, Análisis no lineal, *y* Temas especiales. *Los dos últimos capítulos del volumen I han sido repetidos y son los dos primeros capítulos del volumen II, de modo que toda la materia relativa al álgebra lineal está completa en cada volumen.*

La parte 1 contiene una introducción al álgebra lineal, incluyendo transformaciones lineales, matrices, determinantes, autovalores y formas cuadráticas. Se dan aplicaciones al análisis, en particular al estudio de las ecuaciones diferenciales lineales. Se estudian los sistemas de ecuaciones diferenciales con la ayuda del cálculo matricial. Se demuestran los teoremas de existencia y unicidad por medio del método de Picard de aproximaciones sucesivas, que también se trata utilizando los operadores de contracción.

En la parte 2 se discute el cálculo de funciones de varias variables. El cálculo diferencial se unifica y simplifica con la ayuda del álgebra lineal. Se incluyen reglas de la cadena para campos escalares y vectoriales, y aplicaciones a las ecuaciones diferenciales en derivadas parciales y a problemas de extremos. En cálculo integral se incluyen integrales de línea, integrales múltiples y de superficie, con aplicaciones al análisis vectorial. En esto la exposición sigue más o menos la línea clásica y no incluye un desarrollo formal de las formas diferenciales.

Los temas especiales tratados en la parte 3 son Probabilidades *y* Análisis numérico. *El de probabilidades está dividido en dos capítulos, uno que trata de los espacios muestrales finitos o infinitos numerables; el otro de espacios muestrales no numerables, variables aleatorias, y funciones de distribución. Las aplicaciones se ilustran en el estudio de variables aleatorias uni- y bi-dimensionales.*

El último capítulo contiene una introducción al análisis numérico, poniendo especial atención en los distintos tipos de polinomios de aproximación. Termina el libro con un estudio de las fórmulas de integración aproximada, tales como la regla de Simpson y una discusión de la fórmula de sumación de Euler.

En este volumen hay materia suficiente para un curso anual completo con tres o cuatro sesiones semanales. Presupone un conocimiento del cálculo con una variable como se desarrolla en la mayoría de los cursos del primer año de cálculo. El autor ha imaginado el curso con cuatro sesiones semanales, dos de exposición por parte del profesor y dos para preguntar a los alumnos, empleando aproximadamente diez semanas en cada parte y omitiendo las secciones señaladas con asterisco.

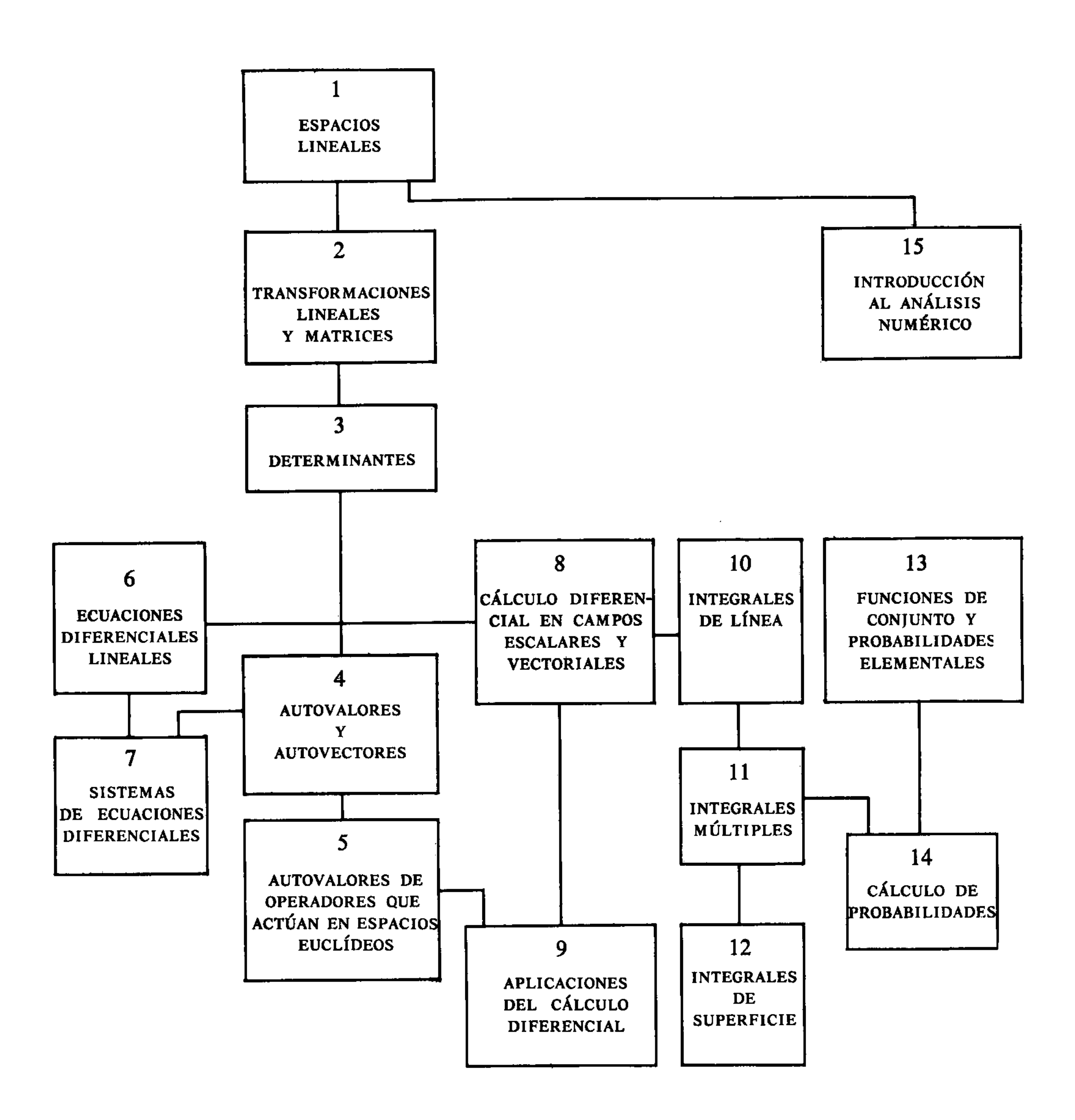
Este segundo volumen ha sido planeado de modo que muchos capítulos pueden omitirse en cursos abreviados. Por ejemplo, el último capítulo de cada parte puede suprimirse sin romper la continuidad de la exposición. La parte primera proporciona material para un curso combinado de álgebra lineal y de ecuaciones diferenciales ordinarias. Cada profesor puede elegir los temas adecuados a sus necesidades y preferencias consultando el diagrama de la página siguiente que muestra la interdependencia lógica de los capítulos.

Una vez más reconozco con agrado el asesoramiento de numerosos amigos y colegas. Al preparar la segunda edición recibí valiosa ayuda de los profesores Herbert S. Zuckerman de la Universidad de Washington, y Basil Gordon de la Universidad de California, Los Ángeles, cada uno de los cuales sugirió varias mejoras. Agradezco también al personal de la Blaisdell Publishing Company su cooperación y ayuda.

Como en otras ocasiones me da especial satisfacción expresar mi gratitud a mi esposa por su valiosa y variada contribución. En reconocimiento le dedico gustosamente este libro.

T. M. A.

Pasadena, California


1
ESPACIOS LINEALES
2
TRANSFORMACIONES LINEALES Y MATRICES
3
DETERMINANTES
4
AUTOVALORES Y AUTOVECTORES
5
AUTOVALORES DE OPERADORES QUE ACTÚAN EN ESPACIOS EUCLÍDEOS
6
ECUACIONES DIFERENCIALES LINEALES
7
SISTEMAS DE ECUACIONES DIFERENCIALES
8
CÁLCULO DIFERENCIAL EN CAMPOS ESCALARES Y VECTORIALES
9
APLICACIONES DEL CÁLCULO DIFERENCIAL
10
INTEGRALES DE LÍNEA
11
INTEGRALES MÚLTIPLES
12
INTEGRALES DE SUPERFICIE
13
FUNCIONES DE CONJUNTO Y PROBABILIDADES ELEMENTALES
14
CÁLCULO DE PROBABILIDADES
15
INTRODUCCIÓN AL ANÁLISIS NUMÉRICO

ÍNDICE ANALÍTICO

Parte 1. Análisis lineal

1. ESPACIOS LINEALES

2. TRANSFORMACIONES LINEALES Y MATRICES

3. DETERMINANTES

4. AUTOVALORES Y AUTOVECTORES

5. AUTOVALORES DE OPERADORES EN ESPACIOS EUCLÍDEOS

6. ECUACIONES DIFERENCIALES LINEALES

7. SISTEMAS DE ECUACIONES DIFERENCIALES

Parte 2. Análisis no lineal

8. CÁLCULO DIFERENCIAL EN CAMPOS ESCALARES Y VECTORIALES

9. APLICACIONES DE CÁLCULO DIFERENCIAL

10. INTEGRALES DE LÍNEA

11. INTEGRALES MÚLTIPLES

12. INTEGRALES DE SUPERFICIE

Parte 3. Temas especiales

13. FUNCIONES DE CONJUNTO Y PROBABILIDAD ELEMENTAL

14. CÁLCULO DE PROBABILIDADES

15. INTRODUCCIÓN AL ANÁLISIS NUMÉRICO

PARTE 1
Análisis lineal

1

ESPACIOS LINEALES

1.1 Introducción

A lo largo de la Matemática se encuentran muchos ejemplos de objetos matemáticos que pueden sumarse unos con otros y multiplicarse por números reales. Ante todo, los números reales son objetos de tal naturaleza. Otros ejemplos son las funciones vectoriales, los números complejos, las series y los vectores en el espacio *n-dimensional*. En este capítulo tratamos un concepto matemático general, llamado *espacio lineal,* que incluye todos esos ejemplos y muchos otros como casos particulares.

Brevemente, un espacio lineal es un conjunto de elementos de naturaleza cualquiera sobre el que pueden realizarse ciertas operaciones llamadas *adición* y *multiplicación por números.* Al definir un espacio lineal no especificamos la naturaleza de los elementos ni decimos cómo se realizan las operaciones entre ellos. En cambio, exigimos que las operaciones tengan ciertas propiedades que tomamos como axiomas de un espacio lineal. Vamos ahora a hacer con detalle una descripción de esos axiomas.

1.2 Definición de espacio lineal

Sea V un conjunto no vacío de objetos, llamados *elementos.* El conjunto V se llama espacio lineal si satisface los diez axiomas siguientes que se enuncian en tres grupos.

Axiomas de clausura

AXIOMA 1. CLAUSURA RESPECTO DE LA ADICIÓN. *A todo par de elementos x e y de V corresponde un elemento único de V llamado suma de x e y, designado por $x + y$.*

AXIOMA 2. CLAUSURA RESPECTO DE LA MULTIPLICACIÓN POR NÚMEROS REALES. *A todo x de V y todo número real a corresponde un elemento de V llamado producto de a por x, designado por ax.*

Axiomas para la adición

AXIOMA 3. LEY CONMUTATIVA. *Para todo x y todo y de V, tenemos $x + y = y + x$.*

AXIOMA 4. LEY ASOCIATIVA. *Cualesquiera que sean x, y, z de V, tenemos $(x + y) + z = x + (y + z)$.*

AXIOMA 5. EXISTENCIA DE ELEMENTO CERO. *Existe un elemento en V, designado con el símbolo O, tal que*

$$x + O = x \qquad \text{para todo } x \text{ de } V:$$

AXIOMA 6. EXISTENCIA DE OPUESTOS. *Para todo x de V, el elemento $(-1)x$ tiene la propiedad*

$$x + (-1)x = O .$$

Axiomas para la multiplicación por números

AXIOMA 7. LEY ASOCIATIVA. *Para todo x de V y todo par de números reales a y b, tenemos*

$$a(bx) = (ab)x .$$

AXIOMA 8. LEY DISTRIBUTIVA PARA LA ADICIÓN EN V. *Para todo x y todo y de V y todo número real a, tenemos*

$$a(x + y) = ax + ay .$$

AXIOMA 9. LEY DISTRIBUTIVA PARA LA ADICIÓN DE NÚMEROS. *Para todo x de V y todo par de números reales a y b, tenemos*

$$(a + b)x = ax + bx .$$

AXIOMA 10. EXISTENCIA DE ELEMENTO IDÉNTICO. *Para todo x de V, tenemos $1x = x$.*

Los espacios lineales así definidos, se llaman, a veces, espacios lineales *reales* para resaltar el hecho de que se multiplican los elementos de V por números reales. Si en los axiomas 2, 7, 8 y 9 se reemplaza *número real* por *número complejo,* la estructura que resulta se llama *espacio lineal complejo.* Algunas veces un espacio lineal se llama también *espacio vectorial lineal* o simplemente *espacio vectorial*; los números utilizados como multiplicadores se llaman *escalares.* Un espacio lineal real tiene números reales como escalares; un espacio lineal complejo tiene como escalares números complejos. Si bien consideraremos principalmente ejemplos de espacios lineales reales, todos los teoremas son válidos para espacios lineales complejos. Cuando digamos espacio lineal sin más, se sobrentenderá que el espacio puede ser real o complejo.

1.3 Ejemplos de espacios lineales

Si precisamos el conjunto V y decimos cómo se suman sus elementos y cómo se multiplican por números, obtenemos un ejemplo concreto de espacio lineal. El lector fácilmente puede comprobar que cada uno de los ejemplos siguientes satisface todos los axiomas para un espacio lineal real.

EJEMPLO 1. Sea $V = \mathbf{R}$, el conjunto de todos los números reales, y sean $x + y$ y ax la adición y la multiplicación ordinarias de números reales.

EJEMPLO 2. Sea $V = \mathbf{C}$ el conjunto de todos los números complejos, definimos $x + y$ como la adición ordinaria de números complejos, y ax como la multiplicación del número complejo x por el número real a. Aunque los elementos de V sean números complejos, éste es un espacio lineal real porque los escalares son reales.

EJEMPLO 3. Sea $V = V_n$, el espacio vectorial de todas las n-plas de números reales, con la adición y la multiplicación por escalares definidas en la forma ordinaria en función de los componentes.

EJEMPLO 4. Sea V el conjunto de todos los vectores V_n ortogonales a un vector no nulo dado N. Si $n = 2$, este espacio lineal es una recta que pasa por O con N como vector normal. Si $n = 3$, es un plano que pasa por O con N como vector normal.

Los siguientes ejemplos se llaman *espacios funcionales.* Los elementos de V son funciones vectoriales, con la suma de dos funciones f y g definidas en la forma ordinaria:

$$(f + g)(x) = f(x) + g(x)$$

para todo real x en la intersección de los dominios de f y g. La multiplicación de una función f por un escalar real a se define así: af es aquella función cuyo valor en cada x del dominio de f es $af(x)$. El elemento cero es la función cuyos valores son nulos para todo x. El lector puede comprobar fácilmente que cada uno de los conjuntos siguientes es un espacio funcional.

EJEMPLO 5. El conjunto de todas las funciones definidas en un intervalo dado.

EJEMPLO 6. El conjunto de todos los polinomios.

EJEMPLO 7. El conjunto de todos los polinomios de grado $\leq n$, siendo n fijo. (Siempre que consideremos este conjunto, se sobrentenderá que siempre está incluido el polinomio nulo.) El conjunto de todos los polinomios de grado *igual* a n no es una espacio lineal porque no se satisfacen los axiomas de clausura. Por ejemplo, la suma de dos polinomios de grado n puede no ser de grado n.

EJEMPLO 8. El conjunto de todas las funciones continuas en un intervalo dado. Si el intervalo es $[a, b]$, designamos este espacio con $C(a, b)$.

EJEMPLO 9. El conjunto de todas las funciones derivables en un punto dado.

EJEMPLO 10. El conjunto de todas las funciones integrables en un intervalo dado.

EJEMPLO 11. El conjunto de todas las funciones f definidas en el punto 1 siendo $f(1) = 0$. El número 0 es esencial en este ejemplo. Si reemplazamos 0 por un número no nulo c, violamos el axioma de clausura.

EJEMPLO 12. El conjunto de todas las soluciones de una ecuación diferencial lineal homogénea $y'' + ay' + by = 0$, donde a y b son constantes dadas. También aquí es esencial el 0. El conjunto de soluciones de una ecuación diferencial no homogénea no satisface los axiomas de clausura.

Estos ejemplos y muchos otros hacen patente cómo el concepto de espacio lineal está extendido por el Álgebra, la Geometría y el Análisis. Cuando se deduce un teorema de los axiomas de un espacio lineal, obtenemos un resultado válido para cada ejemplo concreto. Unificando varios ejemplos de este modo, conseguimos un conocimiento más profundo en cada uno. En ocasiones el conocimiento de un determinado ejemplo ayuda para anticipar o interpretar resultados válidos para otros ejemplos y pone en evidencia relaciones que de otro modo podrían pasar inadvertidas.

1.4 Consecuencias elementales de los axiomas

Los teoremas que siguen se deducen fácilmente de los axiomas de un espacio lineal.

TEOREMA 1.1. UNICIDAD DEL ELEMENTO CERO. *En cualquier espacio lineal existe un elemento cero y sólo uno.*

Demostración. El axioma 5 nos asegura que existe por lo menos un elemento cero. Supongamos que existan dos, sean O_1 y O_2. Haciendo $x = O_1$ y $O = O_2$ en el axioma 5, obtenemos $O_1 + O_2 = O_1$. Análogamente, haciendo $x = O_2$ y $O = O_1$, encontramos $O_2 + O_1 = O_2$. Pero $O_1 + O_2 = O_2 + O_1$ por la ley conmutativa, así que $O_1 = O_2$.

TEOREMA 1.2. UNICIDAD DE ELEMENTOS OPUESTOS. *En cualquier espacio lineal todo elemento tiene exactamente un opuesto. Esto es, para todo x existe un y, y sólo uno tal que $x + y = O$.*

Demostración. El axioma 6 nos dice que cada x tiene por lo menos un opuesto, a saber $(-1)x$. Supongamos que x tenga dos opuestos, sean y_1 e y_2. Entonces $x + y_1 = O$ y $x + y_2 = O$. Sumando y_2 a los dos miembros de la primera igualdad y aplicando los axiomas 5, 4 y 3, obtenemos que

$$y_2 + (x + y_1) = y_2 + O = y_2 ,$$

y

$$y_2 + (x + y_1) = (y_2 + x) + y_1 = O + y_1 = y_1 + O = y_1 .$$

Por consiguiente $y_1 = y_2$, con lo que x tiene exactamente un opuesto, el elemento $(-1)x$.

Notación. El opuesto de x se designa por $-x$. La diferencia $y - x$ se define como la suma $y + (-x)$.

El teorema siguiente muestra un conjunto de propiedades que rigen los cálculos algebraicos elementales en un espacio lineal.

TEOREMA 1.3. *En un espacio lineal, designemos con x e y dos elementos cualesquiera y con a y b dos escalares cualesquiera. Tenemos entonces las propiedades siguientes:*

a) $0x = O$.
b) $aO = O$.

c) $(-a)x = -(ax) = a(-x)$.
d) *Si* $ax = O$, *entonces* $a = 0$ *o* $x = O$, *o los dos.*
e) *Si* $ax = ay$ *y* $a \neq 0$ *entonces* $x = y$.
f) *Si* $ax = bx$ *y* $x \neq O$, *entonces* $a = b$.
g) $-(x + y) = (-x) + (-y) = -x - y$.
h) $x + x = 2x$, $x + x + x = 3x$, *y en general,* $\sum_{i=1}^{n} x = nx$.

Demostraremos a), b) y c) y dejamos como ejercicios las demostraciones de las otras propiedades.

Demostración de a). Sea $z = 0x$. Deseamos demostrar que $z = O$. Sumando z a sí mismo y aplicando el axioma 9, encontramos que

$$z + z = 0x + 0x = (0 + 0)x = 0x = z\,.$$

Sumemos ahora $-z$ a ambos miembros y obtenemos $z = O$.

Demostración de b). Sea $z = aO$, sumar z a sí mismo, y aplicar el axioma 8.

Demostración de c). Sea $z = (-a)x$. Sumando z a ax y aplicando el axioma 9, encontramos que

$$z + ax = (-a)x + ax = (-a + a)x = 0x = O\,,$$

así que z es el opuesto de ax, $z = -(ax)$. Análogamente, si sumamos $a(-x)$ a ax y aplicamos el axioma 8 y la propiedad b), encontramos que $a(-x) = -(ax)$.

1.5 Ejercicios

En los ejercicios del 1 al 28, determinar si cada uno de los conjuntos dados es un espacio lineal real, si la adición y multiplicación por escalares reales está definida en la forma usual. Para aquellos en los que no es así, decir cuáles son los axiomas que no se cumplen. Las funciones de los ejercicios 1 al 17 son reales. En los ejercicios 3, 4 y 5, cada función tiene un dominio que contiene 0 y 1. En los ejercicios 7 al 12, cada dominio contiene todos los números reales.

1. Todas las funciones racionales.
2. Todas las funciones racionales f/g, con el grado de $f \leq$ que el grado de g (incluyendo $f = 0$).
3. Todas las f con $f(0) = f(1)$.
4. Todas las f con $2f(0) = f'(1)$.
5. Todas las f con $f(1) = 1 + f(0)$.
6. Todas las funciones escalonadas definidas en $[0, 1]$.
7. Todas las f en las que $f(x) \to 0$ cuando $x \to +\infty$.
8. Todas las funciones pares.
9. Todas las funciones impares.

10. Todas las funciones acotadas.
11. Todas las funciones crecientes.
12. Todas las funciones con período 2π.
13. Todas las f integrables en $[0, 1]$ con $\int_0^1 f(x)dx = 0$.
14. Todas las f integrables en $[0, 1]$ con $\int_0^1 f(x)dx \geq 0$.
15. Todas las f que satisfacen $f(x) = f(1 - x)$ para todo x.
16. Todos los polinomios de Taylor de grado $\leq n$ para un n fijo (incluyendo el polinomio cero).
17. Todas las soluciones de una ecuación diferencial lineal homogénea de segundo orden $y'' + P(x)y' + Q(x)y = 0$, siendo P y Q funciones dadas, continuas para todo x.
18. Todas las sucesiones reales acotadas.
19. Todas las sucesiones reales convergentes.
20. Todas las series reales convergentes.
21. Todas las series reales absolutamente convergentes.
22. Todos los vectores (x, y, z) de V_3 con $z = 0$.
23. Todos los vectores (x, y, z) de V_3 con $x = 0$ o $y = 0$.
24. Todos los vectores (x, y, z) de V_3 con $y = 5x$.
25. Todos los vectores (x, y, z) de V_3 con $3x + 4y = 1$, $z = 0$.
26. Todos los vectores (x, y, z) de V_3 que son productos de (1, 2, 3) por escalares.
27. Todos los vectores (x, y, z) de V_3 cuyos componentes satisfacen un sistema de tres ecuaciones lineales de la forma

$$a_{11}x + a_{12}y + a_{13}z = 0\,, \qquad a_{21}x + a_{22}y + a_{23}z = 0\,, \qquad a_{31}x + a_{32}y + a_{33}z = 0\,.$$

28. Todos los vectores de V_n que son combinaciones lineales de dos vectores dados A y B.
29. Sea $V = \mathbf{R}^+$, el conjunto de los números reales positivos. Definamos la «suma» de dos elementos x e y de V como su producto $x \cdot y$ (en el sentido ordinario), y definamos la «multiplicación» de un elemento x de V por un escalar c poniendo x^c. Demostrar que V es un espacio lineal real con el elemento cero.
30. a) Demostrar que el axioma 10 puede deducirse de los otros axiomas.
    b) Demostrar que el axioma 10 no puede deducirse de los otros axiomas si el axioma 6 es reemplazado por el axioma 6': Para todo x de V y todo y de V tenemos que $x+y=O$.
31. Sea S el conjunto de todos los pares ordenados (x_1, x_2) de números reales. En cada caso determinar si S es o no un espacio lineal con las operaciones de adición y multiplicación por escalares definidas como se indica. Si el conjunto no es un espacio lineal, indicar cuáles son los axiomas que no se cumplen.
    a) $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$, $\quad a(x_1, x_2) = (ax_1, 0)$.
    b) $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, 0)$, $\quad a(x_1, x_2) = (ax_1, ax_2)$.
    c) $(x_1, x_2) + (y_1, y_2) = (x_1, x_2 + y_2)$, $\quad a(x_1, x_2) = (ax_1, ax_2)$.
    d) $(x_1, x_2) + (y_1, y_2) = (|x_1 + x_2|, |y_1 + y_2|)$, $\quad a(x_1, x_2) = (|ax_1|, |ax_2|)$.
32. Demostrar las partes de la d) a la h) del teorema 1.3.

1.6 Subespacios de un espacio lineal

Dado un espacio lineal V sea S un subconjunto no vacío de V. Si S es también un espacio lineal, entonces S se llama *subespacio* de V. El teorema que sigue

da un sencillo criterio para determinar si un subconjunto de un espacio lineal es o no un subespacio.

TEOREMA 1.4. *Sea S un subconjunto no vacío de un espacio lineal V. Tal subconjunto S es un subespacio si y sólo si satisface los axiomas de clausura.*

Demostración. Si S es un subespacio, satisface todos los axiomas de un espacio lineal, y por tanto, en particular, los axiomas de clausura.

Demostremos ahora que si S satisface los axiomas de clausura, satisface también los otros. Las leyes conmutativa y asociativa para la adición (axiomas 3 y 4) y los axiomas para la multiplicación por escalares (axiomas del 7 al 10) se satisfacen automáticamente en S porque son válidos para todos los elementos de V. Falta comprobar los axiomas 5 y 6, la existencia del elemento cero en S, y la existencia de un opuesto para cada elemento de S.

Sea x un elemento cualquiera de S. (S tiene por lo menos un elemento ya que no es vacío.) Según el axioma 2, ax está en S para todo escalar a. Tomando $a = 0$, resulta que $0x$ está en S. Pero $0x = O$, en virtud del teorema 1.3 a), con lo cual $O \in S$, y se satisface el axioma 5. Tomando $a = -1$, vemos que $(-1)x$ está en S. Pero $x + (-1)x = O$ ya que x y $(-1)x$ están ambos en V, así que el axioma 6 se satisface en S. Por consiguiente S es un subespacio de V.

DEFINICIÓN. *Sea S un subconjunto no vacío de un espacio lineal V. Un elemento x de V de la forma*

$$x = \sum_{i=1}^{k} c_i x_i ,$$

en donde $x_1, \ldots, x_k$ pertenecen todos a S y $c_1, \ldots, c_k$ son escalares, se denomina combinación lineal de elementos de S. El conjunto de todas las combinaciones lineales finitas de elementos de S satisface los axiomas de clausura y por tanto es un subespacio de V. Decimos de ese subespacio que está generado por S, o también le llamamos la envolvente lineal de S, y lo designamos por L(S). Si S es vacío, definimos L(S) como $\{O\}$, el conjunto consta sólo del elemento cero.

Conjuntos distintos pueden generar el mismo subespacio. Por ejemplo, el espacio V_2 está generado por cada uno de los siguientes conjuntos de vectores: $\{i, j\}$, $\{i, j, i + j\}$, $\{O, i, -i, j, -j, i + j\}$. El espacio de todos los polinomios n $p(t)$ de grado $\leq n$ está generado por el conjunto de $n + 1$ polinomios $\{1, t, t^2, \ldots, t^n\}$. También está generado por el conjunto $\{1, t/2, t^2/3, \ldots, t^n/(n + 1)\}$ y por $\{1, (1 + t), (1 + t)^2, \ldots, (1 + t)^n\}$. El espacio de todos los polinomios está generado por el conjunto infinito de los polinomios $\{1, t, t^2, \ldots\}$.

Al llegar aquí surgen de modo natural numerosas preguntas. Por ejemplo, ¿qué espacios pueden generarse por un número finito de elementos? Si un espacio está generado por un número finito de elementos, ¿cuál es el menor número de elementos necesarios? Para discutir estas cuestiones y otras con ellas relacionadas

introducimos los conceptos de *dependencia, independencia, bases* y *dimensión.* Ya en el volumen I encontramos esas ideas al estudiar el espacio vectorial V_n Ahora vamos a extenderlas a espacios lineales de tipo general.

1.7 Conjuntos dependientes e independientes en un espacio lineal

DEFINICIÓN. *Un conjunto S de elementos de un espacio lineal V se llama dependiente si existe un conjunto finito de elementos distintos de S,* $x_1, \ldots, x_k$, *y un correspondiente conjunto de escalares* $c_1, \ldots, c_k$, *no todos cero, tales que*

$$\sum_{i=1}^{k} c_i x_i = O .$$

El conjunto S se llama independiente si no es dependiente. En tal caso, cualesquiera que sean los elementos distintos $x_1, \ldots, x_k$ *de S y los escalares* $c_1, \ldots, c_k$,

$$\sum_{i=1}^{k} c_i x_i = O \qquad \textit{implica} \quad c_1 = c_2 = \cdots = c_k = 0 .$$

Si bien la dependencia y la independencia son propiedades de los conjuntos de elementos, podemos también aplicar esas denominaciones a los elementos mismos. Por ejemplo, los elementos de un conjunto independiente se llaman elementos independientes.

Si S es un conjunto finito, la definición anterior está de acuerdo con la dada en el Volumen I para el espacio V_n. No obstante, la definición dada aquí no está restringida a conjuntos finitos.

EJEMPLO 1. Si un subconjunto T de un conjunto S es dependiente, el mismo S es dependiente. Esto es lógicamente equivalente a la afirmación de que todo subconjunto de un conjunto independiente es independiente.

EJEMPLO 2. Si un elemento de S es el producto de otro por un escalar, S es dependiente.

EJEMPLO 3. Si $O \in S$. entonces S es dependiente

EJEMPLO 4. El conjunto vacío es independiente.

En el Volumen I fueron discutidos muchos ejemplos de conjuntos dependientes e independientes. Los ejemplos que a continuación se comentan, ilustran esos conceptos en espacios funcionales. En cada caso el espacio lineal fundamental V es el conjunto de todas las funciones reales definidas en la recta real.

EJEMPLO 5. Sean $u_1(t) = \cos^2 t$, $u_2(t) = \text{sen}^2 t$, $u_3(t) = 1$ para todo número real t. La identidad pitagórica prueba que $u_1 + u_2 - u_3 = O$, así que las tres funciones u_1, u_2, u_3 son dependientes.

EJEMPLO 6. Sea $u_k(t) = t^k$ para $k = 0, 1, 2, \ldots,$ y t real. El conjunto $S = \{u_0, u_1, u_2, \ldots\}$ es independiente. Para demostrar esto, basta demostrar que para cada n los $n + 1$ polinomios $u_0, u_1, \ldots, u_n$ son independientes. Una relación de la forma $\sum c_k u_k = O$ significa que

$$\sum_{k=0}^{n} c_k t^k = 0 \tag{1.1}$$

para todo real t. Cuando $t = 0$, encontramos que $c_0 = 0$. Repitiendo el proceso, encontramos que cada coeficiente c_k es cero.

EJEMPLO 7. Si $a_1, \ldots, a_n$ son números reales distintos, las n funciones exponenciales

$$u_1(x) = e^{a_1 x}, \ldots, u_n(x) = e^{a_n x}$$

son independientes. Podemos demostrar esto por inducción sobre n. El resultado es trivial cuando $n = 1$. Por consiguiente, supongamos que es válida para $n - 1$ funciones exponenciales y consideremos los escalares $c_1, \ldots, c_n$ tales que

$$\sum_{k=1}^{n} c_k e^{a_k x} = 0. \tag{1.2}$$

Sea a_M el mayor de los n números $a_1, \ldots, a_n$. Multiplicando ambos miembros de (1.2) por $e^{-a_M x}$, obtenemos

$$\sum_{k=1}^{n} c_k e^{(a_k - a_M)x} = 0. \tag{1.3}$$

Si $k \neq M$, el número $a_k - a_M$ es negativo. Por consiguiente, cuando $x \to +\infty$ en la ecuación (1.3), cada término con $k \neq M$ tiende a cero y encontramos que $c_M = 0$. Suprimiendo el término M-ésimo de (1.2) y aplicando la hipótesis de inducción, encontramos que cada uno de los $n - 1$ restantes coeficientes c_k es cero.

TEOREMA 1.5. *Sea $S = \{x_1, \ldots, x_k\}$ un conjunto independiente que consta de k elementos de un espacio lineal V y sea $L(S)$ el subespacio generado por S. Entonces todo conjunto de $k+1$ elementos de $L(S)$ es dependiente.*

Demostración. La demostración es por inducción sobre k, número de elementos de S. Supongamos primero que $k=1$. Entonces, por hipótesis, S contiene un solo elemento x_1 siendo $x_1 \neq O$ puesto que S es independiente. Ahora tomemos en $L(S)$ dos elementos distintos y_1 e y_2. Entonces, cada uno de estos elementos es un escalar multiplicado por x_1, sea $y_1=c_1x_1$ e $y_2=c_2x_1$, en donde c_1 y c_2 no son ambos cero. Multiplicando y_1 por c_2 e y_2 por c_1 y restando, obtenemos

$$c_2y_1 - c_1y_2 = O.$$

Por lo tanto y_1 e y_2 son dependientes, quedando así demostrado el teorema cuando $k=1$.

Supongamos ahora que el teorema es cierto para $k-1$ y demostremos que también lo es para k. Tomemos un conjunto de $k+1$ elementos en $L(S)$, sea $T=\{y_1, y_2, \ldots, y_{k+1}\}$. Queremos probar que T es dependiente. Puesto que cada elemento y_i está contenido en $L(S)$, podemos escribir

$$y_i = \sum_{j=1}^{k} a_{ij}x_j \tag{1.4}$$

para cada $i=1, 2, \ldots, k+1$. Examinemos todos los escalares a_{i1} que multiplican a x_1 y, para ello, consideremos dos casos en la demostración.

CASO 1. $a_{i1}=0$ *para todo* $i=1, 2, \ldots, k+1$. En este caso la suma (1.4) no incluye a x_1, así cada y_i en T está en la envolvente lineal del conjunto $S'=\{x_2, \ldots, x_k\}$. Pero S' es independiente y contiene $k-1$ elementos. Por inducción y para $k-1$, el teorema es cierto, siendo por lo tanto, T dependiente. Esto demuestra el Caso 1.

CASO 2. No son cero todos los escalares a_{i1}. Suponemos que $a_{11} \neq 0$. Tomando $i=1$ en la ecuación (1.4) y multiplicando los dos miembros por c_i, siendo $c_i=a_{i1}/a_{11}$, obtenemos:

$$c_iy_1 = a_{i1}x_1 + \sum_{j=2}^{k} c_ia_{1j}x_j.$$

Si de esta ecuación restamos la (1.4), resulta:

$$c_iy_1 - y_i = \sum_{j=2}^{k} (c_ia_{1j} - a_{ij})x_j,$$

para $i=2, \ldots, k+1$. Esta ecuación expresa cada uno de los elementos $c_iy_1-y_i$ como una combinación lineal de los $k-1$ elementos independientes $x_2, \ldots, x_k$.

Por inducción, los k elementos $c_i y_1 - y_i$ deben ser dependientes. En consecuencia, para cualquier elección de escalares $t_2, \ldots, t_{k+1}$, no todos cero, tenemos

$$\sum_{i=2}^{k+1} t_i(c_i y_1 - y_i) = O,$$

y de aquí deducimos

$$\left(\sum_{i=2}^{k+1} t_i c_i\right) y_1 - \sum_{i=2}^{k+1} t_i y_i = O.$$

Esta es una combinación de $y_1, \ldots, y_{k+1}$, que representa el vector cero, de esta manera los elementos $y_1, \ldots, y_{k+1}$ deben ser dependientes, completando así la demostración.

1.8 Bases y dimensión

DEFINICIÓN. *Un conjunto finito S de elementos de un espacio lineal V se llama base finita de V si S es independiente y genera V. El espacio V es de dimensión finita si tiene una base finita. De otro modo, V es de infinitas dimensiones.*

TEOREMA 1.6. *Sea V un espacio lineal de dimensión finita. Entonces toda base finita de V tiene el mismo número de elementos.*

Demostración. Sean S y T dos bases finitas de V. Supongamos que S y T constan respectivamente de k y m elementos. Puesto que S es independiente y engendra V, el teorema 1.5 nos dice que todo conjunto de $k + 1$ elementos de V es dependiente. Por consiguiente, todo conjunto de más de k elementos de V es dependiente. Ya que T es un conjunto independiente, debe ser $m \leq k$. El mismo razonamiento con S y T intercambiadas prueba que $k \leq m$. Por lo tanto $k = m$.

DEFINICIÓN. *Si un espacio lineal V tiene una base de n elementos, el entero n se llama dimensión de V. Escribimos $n = \dim V$.*

EJEMPLO 1. El espacio V_n tiene dimensión n. Una base es el conjunto de los n vectores coordenados unitarios.

EJEMPLO 2. El espacio de todos los polinomios $p(t)$ de grado $\leq n$ tiene dimensión $n + 1$. Una base es el conjunto de $n + 1$ polinomios $\{1, t, t^2, \ldots, t^n\}$. Todo polinomio de grado $\leq n$ es una combinación lineal de esos $n + 1$ polinomios.

EJEMPLO 3. El espacio de las soluciones de la ecuación diferencial $y'' - 2y' - 3y = 0$ tiene dimensión 2. Una base está formada por las dos funciones $u_1(x) = e^{-x}$, $u_2(x) = e^{3x}$. Toda solución es una combinación lineal de esas dos.

EJEMPLO 4. El espacio de todos los polinomios $p(t)$ es de infinitas dimensiones. El conjunto infinito $\{1, t, t^2, \ldots\}$ genera este espacio y ningún conjunto *finito* de polinomios genera el espacio.

TEOREMA 1.7. *Sea* V *un espacio lineal de dimensión finita con* $\dim V = n$. *Se tiene:*

a) *Cualquier conjunto de elementos independiente de* V *es un subconjunto de una cierta base para* V.

b) *Cualquier conjunto de* n *elementos independientes es una base para* V.

Demostración. Para demostrar (a), consideremos el conjunto independiente $S = \{x_1, \ldots, x_k\}$ constituido por elementos en V. Si $L(S) = V$, entonces S es una base. Si no, entonces hay algún elemento y en V que no está en $L(S)$. Añadamos ese elemento a S y consideremos el nuevo conjunto $S' = \{x_1, \ldots, x_k, y\}$. Si en este conjunto dependiente multiplicamos sus elementos por escalares $c_1, \ldots, c_{k+1}$ siendo alguno diferente de cero, estableceremos que

$$\sum_{i=1}^{k} c_i x_i + c_{k+1} y = O .$$

Pero $c_{k+1} \neq 0$ puesto que $x_1, \ldots, x_k$ son independientes. De aquí que podamos resolver esta ecuación respecto a y llegando a la conclusión que $y \in L(S)$, lo que contradice el supuesto de que y no pertenece a $L(S)$. Por lo tanto el conjunto S' es independiente y contiene $k+1$ elementos. Si $L(S') = V$, entonces S' es una base y, siendo S un subconjunto de S', la parte (a) queda demostrada. Si S' no es una base, entonces podemos proceder con S' de igual manera que procedimos con S y considerar otro nuevo conjunto S'' que contiene $k+2$ elementos y es independiente. Si S'' es una base, (a) queda demostrado. Si no, repetimos el proceso. Debemos llegar a una base después de un número finito de etapas, ya que de otra manera obtendríamos un conjunto independiente con $n+1$ elementos, contradiciendo el teorema (1.5). Por eso, la parte (a) del teorema (1.7) queda demostrada.

Para demostrar la parte (b) consideremos un conjunto independiente S con n elementos. Por la parte (a), S es un subconjunto de base B. Pero por el teorema 1.6, la base B tiene exactamente n elementos, por tanto, $S = B$.

1.9 Componentes

Sea V un espacio lineal de dimensión n y consideremos una base cuyos elementos $e_1, \ldots, e_n$ se toman en un cierto orden. Una tal base ordenada la consideramos como una n-pla $(e_1, \ldots, e_n)$. Si $x \in V$, podemos expresar x como una combinación lineal de esos elementos base:

(1.5)
$$x = \sum_{i=1}^{n} c_i e_i .$$

Los coeficientes en esta ecuación determinan una n-pla de números $(c_1, \ldots, c_n)$ que está unívocamente determinada por x. En efecto, si tenemos otra representación de x como combinación lineal de $e_1, \ldots, e_n$, por ejemplo $x = \sum_{i=1}^{n} d_i e_i$, restando de (1.5) encontramos que $\sum_{i=1}^{n} (c_i - d_i)e_i = O$. Pero ya que los elementos base son independientes, eso implica que $c_i = d_i$ para cada i, con lo cual $(c_1, \ldots, c_n) = (d_1, \ldots, d_n)$.

Los componentes de la n-pla ordenada $(c_1, \ldots, c_n)$ determinada por (1.5) se llaman *componentes de x respecto a la base ordenada* $(e_1, \ldots, e_n)$.

1.10 Ejercicios

En cada uno de los ejercicios del 1 al 10, S es el conjunto de todos los vectores (x, y, z) de V_3 cuyos componentes satisfacen la condición que se da. Determinar si S es un subespacio de V_3. Si lo es, calcular $\dim S$.

1. $x = 0$.
2. $x + y = 0$.
3. $x + y + z = 0$.
4. $x = y$.
5. $x = y = z$.
6. $x = y$ o $x = z$.
7. $x^2 - y^2 = 0$.
8. $x + y = 1$.
9. $y = 2x$ y $z = 3x$.
10. $x + y + z = 0$ y $x - y - z = 0$.

Sea P_n el espacio lineal de todos los polinomios de grado $\leq n$, siendo n fijo. En cada ejercicio del 11 al 20, sea S el conjunto de todos los polinomios f de P_n que satisfacen la condición dada. Determinar si S es un subespacio de P_n. Si lo es, calcular $\dim S$.

11. $f(0) = 0$.
12. $f'(0) = 0$.
13. $f''(0) = 0$.
14. $f(0) + f'(0) = 0$.
15. $f(0) = f(1)$.
16. $f(0) = f(2)$.
17. f es par.
18. f es impar.
19. f es de grado $\leq k$, siendo $k < n$, o $f = 0$.
20. f es de grado k, siendo $k < n$, o $f = 0$.
21. En el espacio lineal de todos los polinomios reales $p(t)$, describir el subespacio engendrado por cada uno de los siguientes conjuntos de polinomios y determinar su dimensión.
    a) $\{1, t^2, t^4\}$; b) $\{t, t^3, t^5\}$; c) $\{t, t^2\}$; d) $\{1 + t, (1 + t)^2\}$.
22. En este ejercicio, $L(S)$ es el subespacio generado por un subconjunto S de un espacio lineal V. Demostrar las proposiciones de la a) a la f).
    a) $S \subseteq L(S)$.
    b) Si $S \subseteq T \subseteq V$ y si T es un subespacio de V, entonces $L(S) \subseteq T$. Esta propiedad se expresa diciendo que $L(S)$ es el *menor* subespacio de V que contiene S.
    c) Un subconjunto S de V es un subespacio de V si y sólo si $L(S) = S$.
    d) Si $S \subseteq T \subseteq V$, entonces $L(S) \subseteq L(T)$.
    e) Si S y T son subespacios de V, también lo es $S \cap T$.
    f) Si S y T son subconjuntos de V, entonces $L(S \cap T) \subseteq L(S) \cap L(T)$.
    g) Dar un ejemplo en el que $L(S \cap T) \neq L(S) \cap L(T)$.
23. Sea V el espacio lineal de todas las funciones reales definidas en la recta real. Determinar si cada uno de los siguientes subconjuntos de V es dependiente o independiente. Calcular la dimensión del subespacio generado por cada conjunto.

a) $\{1, e^{ax}, e^{bx}\}$, $a \neq b$.
b) $\{e^{ax}, xe^{ax}\}$.
c) $\{1, e^{ax}, xe^{ax}\}$.
d) $\{e^{ax}, xe^{ax}, x^2e^{ax}\}$.
e) $\{e^x, e^{-x}, \cosh x\}$.
f) $\{\cos x, \operatorname{sen} x\}$.
g) $\{\cos^2 x, \operatorname{sen}^2 x\}$.
h) $\{1, \cos 2x, \operatorname{sen}^2 x\}$.
i) $\{\operatorname{sen} x, \operatorname{sen} 2x\}$.
j) $\{e^x \cos x, e^{-x} \operatorname{sen} x\}$.

24. Sean V un espacio lineal de dimensión finita, y S un subespacio de V. Demostrar cada una de las proposiciones siguientes.
a) S es de dimensión finita y $\dim S \leq \dim V$.
b) $\dim S = \dim V$ si y sólo si $S = V$.
c) Toda base de S es parte de una base de V.
d) Una base de V no contiene necesariamente una base de S.

1.11 Productos interiores, espacios euclídeos. Normas

En la Geometría euclídea ordinaria, aquellas propiedades que cuentan con la posibilidad de medir longitudes de segmentos rectilíneos y ángulos formados por rectas se llaman propiedades *métricas*. En nuestro estudio de V_n, definimos las longitudes y los ángulos en función del producto escalar. Queremos ahora extender esas ideas a espacios lineales más generales. Primero introduciremos una generalización del producto escalar, que llamaremos *producto interior*, y luego definiremos la longitud y el ángulo en función de este producto interior.

El producto escalar $x \cdot y$ de dos vectores $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots y_n)$ de V_n se definió en el Volumen I por la fórmula

$$x \cdot y = \sum_{i=1}^{n} x_i y_i . \tag{1.6}$$

En un espacio lineal general, escribimos (x, y) en lugar de $x \cdot y$ para los productos interiores, y definimos el producto axiomáticamente y no mediante una fórmula. Esto es, establecemos unas ciertas propiedades que queremos que satisfagan los productos interiores y las consideramos como *axiomas*.

DEFINICIÓN. *Un espacio lineal real V tiene un producto interior si a cada par de elementos x e y de V corresponde un número real único (x, y) que satisface los siguientes axiomas cualesquiera que sean x, y, z de V y para todos los escalares reales c.*

1) $(x, y) = (y, x)$ (*conmutatividad, o simetría*).
2) $(x, y + z) = (x, y) + (x, z)$ (*distributividad, o linealidad*).
3) $c(x, y) = (cx, y)$ (*asociatividad, u homogeneidad*).
4) $(x, x) > 0$ *si* $x \neq O$ (*positividad*).

Un espacio lineal con un producto interior se llama *espacio real euclídeo*.

Observación: Haciendo $c = 0$ en (3), encontramos que $(O, y) = 0$ para todo y.

En un espacio lineal complejo, un producto interior (x, y) es un número complejo que satisface los mismos axiomas que los del producto interior real, excepto el de la simetría que se reemplaza por la relación

$$(1') \qquad (x, y) = \overline{(y, x)}, \qquad \text{(\textit{Simetría hermitiana}†)}$$

siendo $\overline{(y, x)}$ el complejo conjugado de (y, x). En el axioma de homogeneidad, el multiplicador escalar c puede ser cualquier número complejo. Del axioma de la homogeneidad y (1'), llegamos a la relación

$$(3') \qquad (x, cy) = \overline{(cy, x)} = \bar{c}\overline{(y, x)} = \bar{c}(x, y).$$

Un espacio lineal complejo con un producto interior se llama *espacio euclídeo complejo.* (A veces se usa también la denominación de *espacio unitario.*) Un ejemplo es el espacio vectorial complejo $V_n(\mathbf{C})$ brevemente discutido en la sección 12.16 del Volumen I.

Aunque nos interesan principalmente los ejemplos de espacios euclídeos reales, los teoremas de este capítulo son válidos para espacios euclídeos complejos. Cuando decimos espacio euclídeo, sin más, entenderemos que puede ser real o complejo.

El lector debiera comprobar que cada ejemplo que sigue satisface todos los axiomas del producto interior.

EJEMPLO 1. En V_n sea $(x, y) = x \cdot y$, el producto escalar ordinario de x e y.

EJEMPLO 2. Si $x = (x_1, x_2)$ e $y = (y_1, y_2)$ son dos vectores de V_2, definimos (x, y) mediante la fórmula

$$(x, y) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2.$$

Este ejemplo pone de manifiesto que pueden existir más de un producto interior en un espacio lineal dado.

EJEMPLO 3. Sea $C(a, b)$ el espacio lineal de todas las funciones reales con-

† En honor de Charles Hermite (1822-1901) matemático francés que contribuyó mucho al desarrollo del álgebra y del análisis.

tinuas en un intervalo $[a, b]$. Definamos un producto interior de dos funciones f y g con la fórmula

$$(f, g) = \int_a^b f(t)g(t)\, dt .$$

Esta fórmula es análoga a la ecuación (1.6). que define el producto escalar de dos vectores en V_n. Los valores de las funciones $f(t)$ y $g(t)$ desempeñan el papel de los componentes x_i e y_i, y la integración el de la suma.

EJEMPLO 4. En el espacio $C(a, b)$, definimos

$$(f, g) = \int_a^b w(t)f(t)g(t)\, dt ,$$

donde w es una función positiva fija de $C(a, b)$. Tal función se llama *función peso.* En el ejemplo 3 tenemos $w(t) = 1$ para todo t.

EJEMPLO 5. En el espacio lineal de todos los polinomios reales, definimos

$$(f, g) = \int_0^\infty e^{-t}f(t)g(t)\, dt .$$

Debido al factor exponencial, esta integral impropia converge para todo par de polinomios f y g.

TEOREMA 1.8. *En un espacio euclídeo V, todo producto interior satisface la desigualdad de Cauchy-Schwarz:*

$$|(x, y)|^2 \leq (x, x)(y, y) \qquad \textit{para todo } x \textit{ y todo } y \textit{ en } V .$$

Además, el signo de igualdad es válido si y sólo si x e y son dependientes.

Demostración. Si ocurre que o bien $x=0$ o $y=0$ la demostración es trivial. Supongamos que x e y no son ambas cero. Sea $z=ax+by$ en donde a y b son escalares que especificaremos después. Tenemos la desigualdad $(z,z) \geq 0$ para todo a y b. Cuando expresamos esta desigualdad en función de x e y con una elección apropiada de a y b, obtenemos la desigualdad de Cauchy-Schwarz.

Para expresar (z,z) en función de x e y usaremos las propiedades (1'). (2) y (3'), obteniendo

$$\begin{aligned}(z, z) &= (ax + by, ax + by) = (ax, ax) + (ax, by) + (by, ax) + (by, by) \\ &= a\bar{a}(x, x) + a\bar{b}(x, y) + b\bar{a}(y, x) + b\bar{b}(y, y) \geq 0 .\end{aligned}$$

Tomando $a=(y,y)$ y suprimiendo en la desigualdad el factor positivo (y,y), resulta

$$(y, y)(x, x) + \bar{b}(x, y) + b(y, x) + b\bar{b} \geq 0.$$

Ahora, hagamos $b=-(x,y)$. Entonces, $\bar{b}=-(y,x)$ y la última desigualdad, una vez simplificada, toma la forma

$$(y, y)(x, x) \geq (x, y)(y, x) = |(x, y)|^2.$$

Esto demuestra la desigualdad de Cauchy-Schwarz. El signo de igualdad es válido si y sólo si $z=O$. Esto ocurre si y sólo si x e y son dependientes.

EJEMPLO. Aplicando el teorema 1.8 al espacio $C(a, b)$ con el producto interior $(f, g) = \int_a^b f(t)g(t)\, dt$, encontramos que la desigualdad de Cauchy-Schwarz se transforma en

$$\left(\int_a^b f(t)g(t)\, dt\right)^2 \leq \left(\int_a^b f^2(t)\, dt\right)\left(\int_a^b g^2(t)\, dt\right).$$

El producto interior puede utilizarse para introducir el concepto métrico de longitud en cualquier espacio euclídeo.

DEFINICIÓN. *En un espacio euclídeo V, el número no negativo $\|x\|$ definido por la ecuación*

$$\|x\| = (x, x)^{1/2}$$

se denomina norma del elemento x.

Cuando la desigualdad de Cauchy-Schwarz se expresa en función de las normas, toma la forma

$$|(x, y)| \leq \|x\|\, \|y\|.$$

Puesto que es posible definir un producto interior de muchas maneras, la norma de un elemento dependerá del producto interior elegido. Esta falta de unicidad era de esperar. Este hecho es análogo al de que podemos asignar números distintos a la medida de la longitud de un segmento rectilíneo dado, según la elección de escala o unidad de medida. El teorema que sigue da las propiedades fundamentales de las normas que no dependen de la elección de producto interior.

TEOREMA 1.9. *En un espacio euclídeo, toda norma tiene las propiedades siguientes para todos los elementos x e y, y todos los escalares c:*

a) $\|x\| = 0$ si $x = O$.

b) $\|x\| > 0$ si $x \neq O$ (*positividad*).

c) $\|cx\| = |c| \, \|x\|$ (*homogeneidad*).

d) $\|x + y\| \leq \|x\| + \|y\|$ (*desigualdad triangular*).

El signo de igualdad es válido en la desigualdad triangular si y sólo si x e y son dependientes.

Demostración. Las propiedades a), b) y c) se deducen inmediatamente de los axiomas del producto interior. Para demostrar d) observemos que

$$\|x + y\|^2 = (x + y, x + y) = (x, x) + (y, y) + (x, y) + (y, x) =$$
$$= \|x\|^2 + \|y\|^2 + (x, y) + \overline{(x, y)} .$$

La suma $(x, y) + \overline{(x, y)}$ es real. La desigualdad de Cauchy-Schwarz prueba que $|(x, y)| \leq \|x\| \, \|y\|$ y que $|\overline{(x, y)}| \leq \|x\| \, \|y\|$, así que tenemos

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\| \, \|y\| = (\|x\| + \|y\|)^2 .$$

Esto demuestra d). El signo de igualdad en d) es válido siempre que lo sea en la desigualdad de Cauchy-Schwarz. Cuando $y = cx$, siendo $c > 0$, tenemos

$$\|x + y\| = \|x + cx\| = (1 + c) \|x\| = \|x\| + \|cx\| = \|x\| + \|y\| .$$

DEFINICIÓN. *En un espacio euclídeo real V, el ángulo formado por dos elementos no nulos x e y se define como el número θ del intervalo $0 \leq \theta \leq \pi$ que satisface la ecuación*

$$\cos \theta = \frac{(x, y)}{\|x\| \, \|y\|} . \tag{1.7}$$

Observación: La desigualdad de Cauchy-Schwarz prueba que el cociente del segundo miembro de (1.7) está en el intervalo $[-1, 1]$, así que existe sólo un θ en $[0, \pi]$ cuyo coseno es igual al de este cociente.

1.12 Ortogonalidad en un espacio euclídeo

DEFINICIÓN. *En un espacio euclídeo V, dos elementos x e y se llaman ortogonales si su producto interior es cero. Un subconjunto S de V es un conjunto ortogonal si $(x, y) = 0$ para todo par de elementos distintos x e y de S. Un conjunto ortogonal se llama ortonormal si cada uno de sus elementos tiene norma 1.*

El elemento cero es ortogonal a todo elemento de V; es el único elemento ortogonal a sí mismo. El siguiente teorema demuestra una relación entre ortogonalidad y dependencia.

TEOREMA 1.10. *En un espacio euclídeo V, todo conjunto ortogonal de elementos no nulos es independiente. En particular, en un espacio euclídeo de dimensión finita con* $\dim V = n$, *todo conjunto ortogonal que conste de n elementos no nulos es una base para V.*

Demostración. Sea S un conjunto ortogonal de elementos no nulos de V, y supongamos que una cierta combinación lineal finita de elementos de S es cero, sea

$$\sum_{i=1}^{k} c_i x_i = O,$$

donde cada $x_i \in S$. Formando el producto escalar de cada miembro por x_1 y teniendo en cuenta que $(x_1, x_i) = 0$ si $i \neq 1$, encontramos que $c_1(x_1, x_1) = 0$. Pero $(x_1, x_1) \neq 0$ ya que $x_1 \neq O$ con lo cual $c_1 = 0$. Repitiendo el razonamiento cambiando x_1 por x_j, encontramos que cada $c_j = 0$. Esto prueba que S es independiente. Si $\dim V = n$ y si S consta de n elementos, el teorema 1.7 b) demuestra que S es una base para V.

EJEMPLO. En el espacio lineal real $C(0, 2\pi)$ con el producto interior $(f, g) = \int_0^{2\pi} f(x)g(x)\,dx$, sea S el conjunto de las funciones trigonométricas $\{u_0, u_1, u_2, \ldots\}$ dadas por

$$u_0(x) = 1, \qquad u_{2n-1}(x) = \cos nx, \qquad u_{2n}(x) = \operatorname{sen} nx, \qquad \text{para} \quad n = 1, 2, \ldots.$$

Si $m \neq n$, tenemos las relaciones de ortogonalidad

$$\int_0^{2\pi} u_n(x)u_m(x)\,dx = 0,$$

así que S es un conjunto ortogonal. Puesto que ningún elemento de S es el elemento cero, S es independiente. La norma de cada elemento de S se calcula fácilmente. Tenemos $(u_0, u_0) = \int_0^{2\pi} dx = 2\pi$ y, para $n \geq 1$, tenemos

$$(u_{2n-1}, u_{2n-1}) = \int_0^{2\pi} \cos^2 nx\,dx = \pi, \qquad (u_{2n}, u_{2n}) = \int_0^{2\pi} \operatorname{sen}^2 nx\,dx = \pi.$$

Por consiguiente, $\|u_0\| = \sqrt{2\pi}$ y $\|u_n\| = \sqrt{\pi}$ para $n \geq 1$. Dividiendo cada u_n por su norma, obtenemos un conjunto ortonormal $\{\varphi_0, \varphi_1, \varphi_2, \ldots\}$ donde $\varphi_n = u_n/\|u_n\|$. Así pues, tenemos

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}}, \qquad \varphi_{2n-1}(x) = \frac{\cos nx}{\sqrt{\pi}}, \qquad \varphi_{2n}(x) = \frac{\operatorname{sen} nx}{\sqrt{\pi}}, \qquad \text{para} \quad n \geq 1.$$

En la sección 1.14 demostraremos que todo espacio euclídeo de dimensión finita tiene una base ortogonal. El teorema que sigue muestra cómo se calculan los componentes de un elemento relativos a una tal base.

TEOREMA 1.11. *Sea V un espacio euclídeo de dimensión finita n, y supongamos que* $S = \{e_1, \ldots, e_n\}$ *es una base ortogonal para V. Si un elemento x está expresado como una combinación lineal de los elementos de la base, sea ésta*

$$x = \sum_{i=1}^{n} c_i e_i, \tag{1.8}$$

entonces sus componentes relativos a la base ordenada $(e_1, \ldots, e_n)$ *vienen dados por las fórmulas*

$$c_j = \frac{(x, e_j)}{(e_j, e_j)} \quad \textit{para} \quad j = 1, 2, \ldots, n. \tag{1.9}$$

En particular, si S es una base ortonormal, cada c_j *viene dada por*

$$c_j = (x, e_j). \tag{1.10}$$

Demostración. Formando el producto interior de cada miembro de (1,8) con e_j, obtenemos

$$(x, e_j) = \sum_{i=1}^{n} c_i(e_i, e_j) = c_j(e_j, e_j)$$

puesto que $(e_i, e_j) = 0$ si $i \neq j$. Esto implica (1.9), y cuando $(e_j, e_j) = 1$, obtenemos (1.10).

Si $\{e_1, \ldots, e_n\}$ es una base ortonormal, la ecuación (1.9) puede escribirse en la forma

$$x = \sum_{i=1}^{n} (x, e_i) e_i. \tag{1.11}$$

El siguiente teorema prueba que en un espacio euclídeo real de dimensión finita con una base ortonormal el producto interior de dos elementos es igual a la suma de los productos de sus componentes.

TEOREMA 1.12. *Sea V un espacio euclídeo real de dimensión finita n, y supongamos que $\{e_1, \ldots, e_n\}$ es una base ortonormal para V. Para todo par de elementos x e y de V, tenemos*

$$(1.12) \qquad (x, y) = \sum_{i=1}^{n} (x, e_i)\overline{(y, e_i)} \quad (\textit{Fórmula de Parseval}).$$

En particular, cuando $x = y$, tenemos

$$(1.13) \qquad \|x\|^2 = \sum_{i=1}^{n} |(x, e_i)|^2 .$$

Demostración. Formando el producto interior de ambos miembros de la ecuación (1.11) con y, y aplicando la propiedad de linealidad del producto interior, obtenemos (1.12). Cuando $x = y$, la ecuación (1.12) se reduce a (1.13).

Observación: La ecuación (1.12) se denomina como se indica en honor de M. A. Parseval (aproximadamente 1776-1836), que obtuvo este tipo de fórmula en un espacio funcional especial. La ecuación (1.13) es una generalización del teorema de Pitágoras.

1.13 Ejercicios

1. Sean $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ vectores arbitrarios de V_n. Determinar en cada caso, si (x, y) es un producto interior en V_n, si (x, y) está definido por la fórmula que se da. En el caso en que (x, y) no sea un producto interior, decir cuáles son los axiomas que no se satisfacen.

a) $(x, y) = \sum_{i=1}^{n} x_i |y_i|$.

b) $(x, y) = \left| \sum_{i=1}^{n} x_i y_i \right|$.

c) $(x, y) = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j$.

d) $(x, y) = \left(\sum_{i=1}^{n} x_i^2 y_i^2 \right)^{1/2}$.

e) $(x, y) = \sum_{i=1}^{n} (x_i + y_i)^2 - \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2$.

2. Supongamos que mantenemos los tres primeros axiomas del producto interior real (simetría, linealidad y homogeneidad) pero reemplazamos el cuarto axioma por uno nuevo (4'): $(x, x) = 0$ si y sólo si $x = O$. Demostrar que o $(x, x) > 0$ para todo $x \neq O$ o bien $(x, x) < 0$ para todo $x \neq O$.

[*Indicación:* Suponer $(x, x) > 0$ para un cierto $x \neq O$ y $(y, y) < 0$ para un cierto $y \neq O$. En el espacio generado por $\{x, y\}$, hallar un elemento $z \neq O$ con $(z, z) = 0$.]

Demostrar que en los ejercicios del 3 al 7 cada una de las proposiciones es válida para todo par de elementos x e y de un espacio euclídeo real.

3. $(x, y) = 0$ si y sólo si $\|x + y\| = \|x - y\|$.
4. $(x, y) = 0$ si y sólo si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.
5. $(x, y) = 0$ si y sólo si $\|x + cy\| \geq \|x\|$ para todo c real
6. $(x + y, x - y) = 0$ si y sólo si $\|x\| = \|y\|$.
7. Si x e y son elementos no nulos que forman un ángulo θ, entonces

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\,\|x\|\,\|y\| \cos \theta .$$

8. En el espacio lineal real $C(1, e)$, definimos un producto interior por

$$(f, g) = \int_1^e (\log x) f(x) g(x)\, dx .$$

a) Si $f(x) = \sqrt{x}$, calcular $\|f\|$.

b) Hallar un polinomio de primer grado $g(x) = a + bx$ que sea ortogonal a la función constante $f(x) = 1$.

9. En el espacio lineal real $C(-1, 1)$, sea $(f, g) = \int_{-1}^{1} f(t) g(t) dt$. Considerar las tres funciones u_1, u_2 u_3 dadas por

$$u_1(t) = 1 , \qquad u_2(t) = t , \qquad u_3(t) = 1 + t .$$

Demostrar que dos de ellas son ortogonales, dos forman entre sí un ángulo $\pi/3$, y dos forman entre sí un ángulo $\pi/6$.

10. En el espacio lineal P_n de todos los polinomios reales de grado $\leq n$, definimos

$$(f, g) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) g\left(\frac{k}{n}\right) .$$

a) Demostrar que (f, g) es un producto interior para P_n.

b) Calcular (f, g) cuando $f(t) = t$ y $g(t) = at + b$.

c) Si $f(t) = t$, hallar todos los polinomios g ortogonales a f.

11. En el espacio lineal de todos los polinomios reales, definimos $(f, g) = \int_0^\infty e^{-t} f(t) g(t)\, dt$.

a) Demostrar que esa integral impropia converge absolutamente para todos los polinomios f y g.

b) Si $x_n(t) = t^n$ para $n = 0, 1, 2, \ldots$, demostrar que $(x_n, x_m) = (m + n)!$.

c) Calcular (f, g) cuando $f(t) = (t + 1)^2$ y $g(t) = t^2 + 1$.

d) Hallar todos los polinomios de primer grado $g(t) = a + bt$ ortogonales a $f(t) = 1 + t$.

12. En el espacio lineal de todos los polinomios reales, determinar si (f, g) es o no un producto interior cuando se define (f, g) con la fórmula que se da. En el caso en que (f, g) no es un producto interior, indicar qué axiomas no son respetados. En c), f' y g' indican derivadas.

a) $(f,g) = f(1)g(1)$.

b) $(f,g) = \left| \int_0^1 f(t)g(t)\,dt \right|$.

c) $(f,g) = \int_0^1 f'(t)g'(t)\,dt$.

d) $(f,g) = \left(\int_0^1 f(t)\,dt\right)\left(\int_0^1 g(t)\,dt\right)$.

13. V está formado con todas las sucesiones indefinidas de números reales $\{x_n\}$ para los cuales las series $\sum x_n^2$ convergen. Si $x = \{x_n\}$ e $y = \{y_n\}$ son dos elementos de V, definimos

$$(x, y) = \sum_{n=1}^{\infty} x_n y_n .$$

a) Demostrar que esta serie converge absolutamente.

[*Indicación:* Usar la desigualdad de Cauchy-Schwarz para aproximar la suma $\sum_{n=1}^{M} |x_n y_n|$.]

b) Demostrar que V es un espacio lineal con (x, y) como producto interior.

c) Calcular (x, y) si $x_n = 1/n$ e $y_n = 1/(n + 1)$ para $n \geq 1$.

d) Calcular (x, y) si $x_n = 2^n$ e $y_n = 1/n!$ para $n \geq 1$.

14. Sea V el conjunto de todas las funciones reales f continuas en $[0, +\infty)$ y tales que la integral $\int_0^\infty e^{-t} f^2(t)dt$ converge. Definamos $(f, g) = \int_0^\infty e^{-t} f(t)g(t)dt$.

a) Demostrar que la integral que da (f, g) converge absolutamente para cada par de funciones f y g de V.

[*Indicación:* Aplicar la desigualdad de Cauchy-Schwarz para aproximar la integral $\int_0^M e^{-t}|f(t)g(t)|\,dt$.]

b) Demostrar que V es un espacio lineal con (f, g) como producto interior.

c) Calcular (f, g) si $f(t) = e^{-t}$ y $g(t) = t^n$, donde $n = 0, 1, 2, \ldots$.

15. En un espacio euclídeo complejo, demostrar que el producto interior tiene las siguientes propiedades para todos los elementos x, y, z y todos los complejos a y b.

a) $(ax, by) = a\bar{b}(x, y)$.   b) $(x, ay + bz) = \bar{a}(x, y) + \bar{b}(x, z)$.

16. Demostrar que en todo espacio euclídeo son válidas las identidades siguientes.

a) $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + (x, y) + (y, x)$.

b) $\|x + y\|^2 - \|x - y\|^2 = 2(x, y) + 2(y, x)$.

c) $\|x + y\|^2 + \|x - y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$.

17. Demostrar que el espacio de todas las funciones complejas continuas en un intervalo $[a, b]$ se transforma en un espacio unitario si definimos un producto interior por la fórmula

$$(f, g) = \int_a^b w(t) f(t)\overline{g(t)}\,dt ,$$

donde w es una función positiva fija, continua en $[a, b]$.

1.14 Construcción de conjuntos ortogonales. Método de Gram-Schmidt

Todo espacio lineal de dimensión finita tiene una base finita. Si el espacio es euclídeo, podemos construir siempre una base *ortogonal*. Este resultado se dedu-

cirá como consecuencia de un teorema cuya demostración enseña a construir conjuntos ortogonales en cualquier espacio euclídeo, de dimensión finita o de infinitas dimensiones. La construcción se llama *método de Gram-Schmidt*, en memoria de J. P. Gram (1850-1916) y E. Schmidt (1845-1921).

TEOREMA 1.13. TEOREMA DE ORTOGONALIZACIÓN. *Sea $x_1, x_2, \ldots,$ una sucesión finita o indefinida de elementos de un espacio euclídeo V, y designemos con $L(x_1, \ldots, x_k)$ el subespacio generado por los k primeros de esos elementos. Existe una sucesión correspondiente de elementos $y_1, y_2, \ldots,$ de V que tiene las siguientes propiedades para cada entero k:*

a) *El elemento y_k es ortogonal a todo elemento del subespacio $L(y_1, \ldots y_{k-1})$.*

b) *El subespacio generado por $y_1, \ldots, y_k$ es el mismo que el generado por $x_1, \ldots, x_k$:*

$$L(y_1, \ldots, y_k) = L(x_1, \ldots, x_k).$$

c) *La sucesión $y_1, y_2, \ldots,$ es única, salvo factores escalares. Esto es, si $y_1', y_2', \ldots,$ es otra sucesión de elementos de V que satisfacen las propiedades* a) *y* b), *entonces para cada k existe un escalar c_k tal que $y_k' = c_k y_k$.*

Demostración. Construyamos los elementos $y_1, y_2, \ldots,$ por inducción. Para iniciar el proceso, tomamos $y_1 = x_1$. Supongamos ahora que hemos construido $y_1, \ldots, y_r$ de modo que a) y b) se satisfacen cuando $k = r$. Definamos y_{r+1} mediante la ecuación

$$y_{r+1} = x_{r+1} - \sum_{i=1}^{r} a_i y_i, \tag{1.14}$$

donde los escalares $a_1, \ldots, a_r$ tienen que determinarse. Para $j \leq r$, el producto interior de y_{r+1} con y_j viene dado por

$$(y_{r+1}, y_j) = (x_{r+1}, y_j) - \sum_{i=1}^{r} a_i(y_i, y_j) = (x_{r+1}, y_j) - a_j(y_j, y_j),$$

puesto que $(y_i, y_j) = 0$ si $i \neq j$. Si $y_j \neq O$, podemos hacer y_{r+1} ortogonal a y_j tomando

$$a_j = \frac{(x_{r+1}, y_j)}{(y_j, y_j)}. \tag{1.15}$$

Si $y_j = O$, entonces y_{r+1} es ortogonal a y_j para cualquier a_j que se elija, en este caso elegimos $a_j = 0$. Así pues, el elemento y_{r+1} está bien definido y es ortogonal

a cada uno de los anteriores elementos $y_1, \ldots, y_r$. Por consiguiente, es ortogonal a todo elemento del subespacio

$$L(y_1, \ldots, y_r).$$

Esto demuestra a) cuando $k = r + 1$.

Para demostrar b) cuando $k = r + 1$, tenemos que probar que $L(y_1, \ldots, y_{r+1}) = L(x_1, \ldots, x_{r+1})$, dado que $L(y_1, \ldots, y_r) = L(x_1, \ldots, x_r)$. Los r primeros elementos $y_1, \ldots, y_r$ pertenecen a

$$L(x_1, \ldots, x_r)$$

y por tanto están en el subespacio más amplio $L(x_1, \ldots, x_{r+1})$. El nuevo elemento y_{r+1} dado por (1.14) es una diferencia de dos elementos de $L(x_1, \ldots, x_{r+1})$ así que también está en $L(x_1, \ldots, x_{r+1})$. Esto demuestra que

$$L(y_1, \ldots, y_{r+1}) \subseteq L(x_1, \ldots, x_{r+1}).$$

La ecuación (1.14) prueba que x_{r+1} es la suma de dos elementos de $L(y_1, \ldots, y_{r+1})$ con lo que un razonamiento análogo da la inclusión en el otro sentido:

$$L(x_1, \ldots, x_{r+1}) \subseteq L(y_1, \ldots, y_{r+1}).$$

Esto demuestra b) cuando $k = r + 1$. Por lo tanto a) y b) han sido demostrados por inducción respecto de k.

Finalmente demostramos c) por inducción respecto de k. El caso $k = 1$ es trivial. Por consiguiente, supongamos que c) es cierto para $k = r$ y consideremos el elemento y'_{r+1}. En virtud de b), este elemento pertenece a

$$L(y_1, \ldots, y_{r+1}),$$

así que podemos escribir

$$y'_{r+1} = \sum_{i=1}^{r+1} c_i y_i = z_r + c_{r+1} y_{r+1},$$

donde $z_r \in L(y_1, \ldots, y_r)$. Queremos demostrar que $z_r = O$. Por la propiedad a), y'_{r+1} y $c_{r+1}y_{r+1}$ son ambos ortogonales a z_r. Por consiguiente, su diferencia, z_r, es ortogonal a z_r. Dicho de otro modo, z_r es ortogonal a sí mismo, así que $z_r = O$. Esto completa la demostración del teorema de ortogonalidad.

En la construcción anterior, puede suceder que $y_{r+1} = O$ para algún r. Entonces (1.14) prueba que x_{r+1} es una combinación lineal de $y_1, \ldots, y_r$, y por tanto

de $x_1, \ldots, x_r$, así que los elementos $x_1, \ldots, x_{r+1}$ son dependientes. En otras palabras, si los k primeros elementos $x_1, \ldots, x_k$ son independientes, los elementos correspondientes $y_1, \ldots, y_k$ son *no nulos*. En este caso los coeficientes a_i de (1.14) vienen dados por (1.15), y las fórmulas que definen $y_1, \ldots, y_k$ se convierten en

$$(1.16) \qquad y_1 = x_1\,, \qquad y_{r+1} = x_{r+1} - \sum_{i=1}^{r} \frac{(x_{r+1}, y_i)}{(y_i, y_i)}\, y_i \quad \text{para} \quad r = 1, 2, \ldots, k-1\,.$$

Estas fórmulas constituyen el método de Gram-Schmidt para construir un conjunto ortogonal de elementos no nulos $y_1, \ldots, y_k$ que generan el mismo subespacio que el conjunto independiente dado $x_1, \ldots x_k$. En particular, si $x_1, \ldots, x_k$ es una base para un espacio euclídeo de dimensión finita, entonces $y_1, \ldots, y_k$ *es* una base ortogonal para el mismo espacio. También podemos convertir ésta en una base ortonormal *normalizando* cada uno de los elementos y_i, esto es, dividiéndolo por su norma. Por consiguiente, como corolario del teorema 1.13 tenemos el siguiente.

TEOREMA 1.14. *Todo conjunto euclídeo de dimensión finita tiene una base ortonormal.*

Si x e y son elementos en un espacio euclídeo, con $y \neq O$, el elemento

$$\frac{(x, y)}{(y, y)}\, y$$

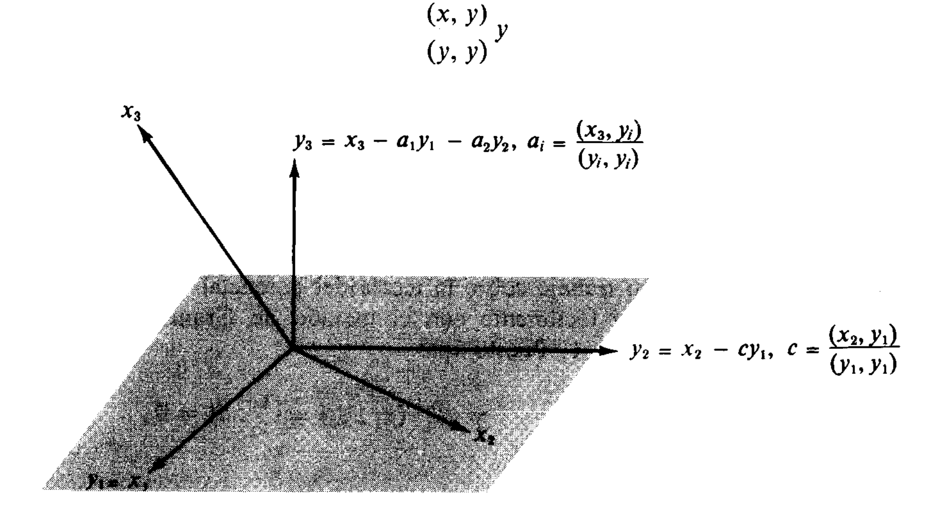


FIGURA 1.1 *El método de Gram-Schmidt en V_3. Un conjunto ortogonal $\{y_1, y_2, y_3\}$ se construye a partir de un conjunto independiente $\{x_1, x_2, x_3\}$.*

se llama la *proyección de x sobre y*. En el método de Gram-Schmidt (1.16), construimos el elemento y_{r+1} restando de x_{r+1} la proyección de x_{r+1} sobre cada uno de los anteriores elementos $y_1, \ldots, y_r$. La figura 1.1 representa la construcción geométrica en el espacio vectorial V_3.

EJEMPLO 1. En V_4, hallar una base ortonormal para el subespacio generado por los tres vectores $x_1 = (1, -1, 1, -1)$, $x_2 = (5, 1, 1, 1,)$, y $x_3 = (-3, -3, 1, -3)$.

Solución. Aplicando el método de Gram-Schmidt, encontramos

$$y_1 = x_1 = (1, -1, 1, -1)\,,$$

$$y_2 = x_2 - \frac{(x_2, y_1)}{(y_1, y_1)}\, y_1 = x_2 - y_1 = (4, 2, 0, 2)\,,$$

$$y_3 = x_3 - \frac{(x_3, y_1)}{(y_1, y_1)}\, y_1 - \frac{(x_3, y_2)}{(y_2, y_2)}\, y_2 = x_3 - y_1 + y_2 = (0, 0, 0, 0)\,.$$

Puesto que $y_3 = O$, los tres vectores x_1, x_2, x_3 deben ser dependientes. Pero ya que y_1 e y_2 son no nulos, los vectores x_1 y x_2 son independientes. Por consiguiente $L(x_1, x_2, x_3)$ es un subespacio de dimensión 2. El conjunto $\{y_1, y_2\}$ es una base ortogonal para ese subespacio. Dividiendo y_1 e y_2 cada uno por su norma llegamos a una base ortonormal que consta de dos vectores

$$\frac{y_1}{\|y_1\|} = \frac{1}{2}(1, -1, 1, -1) \qquad \text{y} \qquad \frac{y_2}{\|y_2\|} = \frac{1}{\sqrt{6}}(2, 1, 0, 1)\,.$$

EJEMPLO 2. *Polinomios de Legendre.* En el espacio lineal de todos los polinomios, con el producto interior $(x, y) = \int_{-1}^{1} x(t)\, y(t)\, dt$, consideramos la sucesión indefinida $x_0, x_1, x_2, \ldots$, donde $x_n(t) = t^n$. Cuando se aplica a esa sucesión el teorema de ortogonalización se transforma en otra sucesión de polinomios $y_0, y_1, y_2, \ldots$, que el matemático francés A. M. Legendre (1752-1833) fue el primero en encontrar en su trabajo sobre la teoría del potencial. Los primeros de esos polinomios se calculan fácilmente con el método de Gram-Schmidt. Ante todo, tenemos $y_0(t) = x_0(t) = 1$. Puesto que

$$(y_0, y_0) = \int_{-1}^{1} dt = 2 \qquad \text{y} \qquad (x_1, y_0) = \int_{-1}^{1} t\, dt = 0\,,$$

encontramos que

$$y_1(t) = x_1(t) - \frac{(x_1, y_0)}{(y_0, y_0)}\, y_0(t) = x_1(t) = t\,.$$

A continuación, utilizamos las relaciones

$$(x_2, y_0) = \int_{-1}^{1} t^2\,dt = \frac{2}{3}, \qquad (x_2, y_1) = \int_{-1}^{1} t^3\,dt = 0, \qquad (y_1, y_1) = \int_{-1}^{1} t^2\,dt = \frac{2}{3}$$

para obtener

$$y_2(t) = x_2(t) - \frac{(x_2, y_0)}{(y_0, y_0)}\, y_0(t) - \frac{(x_2, y_1)}{(y_1, y_1)}\, y_1(t) = t^2 - \frac{1}{3}.$$

Del mismo modo, encontramos que

$$y_3(t) = t^3 - \frac{3}{5}t, \qquad y_4(t) = t^4 - \frac{6}{7}t^2 + \frac{3}{35}, \qquad y_5(t) = t^5 - \frac{10}{9}t^3 + \frac{5}{21}t.$$

En el capítulo 6 encontraremos de nuevo esos polinomios en el estudio de las ecuaciones diferenciales, y probaremos que

$$y_n(t) = \frac{n!}{(2n)!}\frac{d^n}{dt^n}(t^2 - 1)^n.$$

Los polinomios P_n dados por

$$P_n(t) = \frac{(2n)!}{2^n(n!)^2}\, y_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2 - 1)^n$$

se conocen con el nombre de *polinomios de Legendre.* Los polinomios de la sucesión ortonormal correspondiente $\varphi_0, \varphi_1, \varphi_2, \ldots,$ dados por $\varphi_n = y_n/\|y_n\|$ se llaman *polinomios de Legendre normalizados.* De las fórmulas para $y_0, \ldots, y_5$ dadas antes, encontramos que

$$\varphi_0(t) = \frac{1}{\sqrt{2}}, \quad \varphi_1(t) = \sqrt{\frac{3}{2}}\,t, \qquad \varphi_2(t) = \frac{1}{2}\sqrt{\frac{5}{2}}(3t^2 - 1), \quad \varphi_3(t) = \frac{1}{2}\sqrt{\frac{7}{2}}(5t^3 - 3t),$$

$$\varphi_4(t) = \frac{1}{8}\sqrt{\frac{9}{2}}(35t^4 - 30t^2 + 3), \qquad \varphi_5(t) = \frac{1}{8}\sqrt{\frac{11}{2}}(63t^5 - 70t^3 + 15t).$$

1.15 Complementos ortogonales. Proyecciones

Sean V un espacio euclídeo y S un subespacio de dimensión finita. Vamos a considerar el siguiente problema de aproximación: *Dado un elemento x de*

V, determinar un elemento en S cuya distancia a x sea lo más pequeña posible. La distancia entre dos elementos x e y se define como la norma $\|x - y\|$.

Antes de discutir este problema en su forma general, consideremos un caso particular, representado en la figura 1.2. Aquí V es el espacio vectorial V_3 y S es un subespacio de dimensión dos, un plano que pasa por el origen. Dado x de V, el problema consiste en encontrar, en el plano S, el punto s más próximo a x.

Si $x \in S$, evidentemente la solución es $s = x$. Si x no pertenece a S, el punto más próximo s se obtiene trazando una perpendicular desde x al plano. Este sencillo ejemplo sugiere una introducción al problema general de aproximación y da origen a la discusión que sigue.

DEFINICIÓN. *Sea S un subconjunto de un espacio euclídeo V. Se dice que un elemento de V es ortogonal a S si es ortogonal a todo elemento de S. El conjunto de todos los elementos ortogonales a S se designa con $S^{\perp}$ y es el «perpendicular a S».*

Es un ejercicio sencillo comprobar que $S^{\perp}$ es un subespacio de V, tanto, si S lo es como si no lo es. En el caso en que S sea un subespacio, entonces $S^{\perp}$ se llama *complemento ortogonal* de S.

EJEMPLO. Si S es un plano que pasa por el origen, como se ve en la figura 1.2, entonces $S^{\perp}$ es una recta por el origen perpendicular a ese plano. Este ejemplo da también una interpretación geométrica para el teorema siguiente.

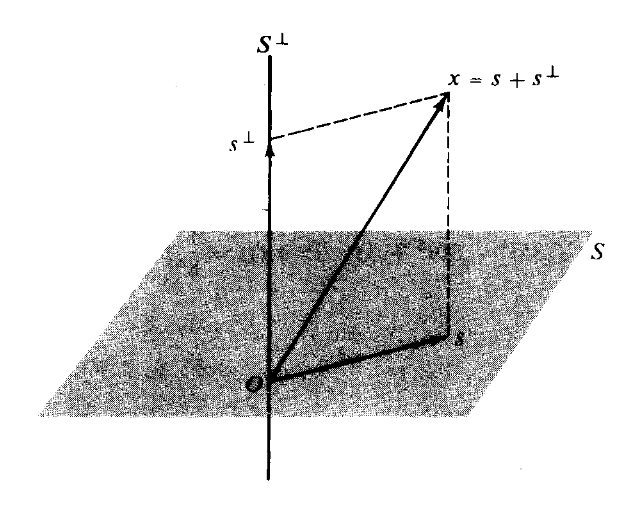


FIGURA 1.2 *Interpretación geométrica del teorema de descomposición ortogonal en V_3.*

TEOREMA 1.15. TEOREMA DE LA DESCOMPOSICION ORTOGONAL. *Sean V un espacio euclídeo y S un subespacio de V de dimensión finita. Todo elemento x de V puede representarse en forma única como una suma de dos elementos, uno de S y otro de $S^\perp$. Esto es, tenemos*

$$x = s + s^\perp, \quad \textit{donde} \quad s \in S \quad y \quad s^\perp \in S^\perp. \tag{1.17}$$

Además, la norma de x viene dada por la fórmula pitagórica

$$\|x\|^2 = \|s\|^2 + \|s^\perp\|^2. \tag{1.18}$$

Demostración. Demostremos primero que existe en realidad una descomposición ortogonal (1.17). Puesto que S es de dimensión finita, tiene una base ortonormal finita, sea ésta $\{e_1, \ldots, e_n\}$. Dado x, definimos los elementos s y $s^\perp$ así:

$$s = \sum_{i=1}^{n} (x, e_i)e_i, \qquad s^\perp = x - s. \tag{1.19}$$

Observemos que cada término $(x, e_i)e_i$ es la proyección de x sobre e_i. El elemento s es la suma de las proyeccciones de x sobre cada elemento de la base. Puesto que s es una combinación lineal de los elementos de la base, s está en S. La definición de $s^\perp$ prueba que la ecuación (1.17) es válida. Para demostrar que $s^\perp$ está en $S^\perp$, consideremos el producto interior de $s^\perp$ y cualquier elemento e_j de la base. Tenemos

$$(s^\perp, e_j) = (x - s, e_j) = (x, e_j) - (s, e_j).$$

Pero de (1.19), encontramos que $(s, e_j) = (x, e_j)$, así que $s^\perp$ es ortogonal a e_j. Por consiguiente $s^\perp$ es ortogonal a todo elemento de S, lo cual significa que $s^\perp \in S^\perp$.

Probamos a continuación que la descomposición ortogonal (1.17) es única. Supongamos que x tuviera dos descomposiciones, sean éstas

$$x = s + s^\perp \qquad \text{y} \qquad x = t + t^\perp, \tag{1.20}$$

donde s y t están en S, y $s^\perp$ y $t^\perp$ están en $S^\perp$. Queremos demostrar que $s = t$ y $s^\perp = t^\perp$. De (1.20), tenemos $s - t = t^\perp - s^\perp$, así que sólo necesitamos demostrar que $s - t = O$. Pero $s - t \in S$ y $t^\perp - s^\perp \in S^\perp$ con lo que $s - t$ es ortogonal a $t^\perp - s^\perp$ e igual a $t^\perp - s^\perp$. Puesto que el elemento cero es el único elemento ortogonal a sí mismo, debe ser $s - t = O$. Esto demuestra que la descomposición es única.

Finalmente, demostremos que la norma de x viene dada por la fórmula pitagórica. Tenemos

$$\|x\|^2 = (x, x) = (s + s^\perp, s + s^\perp) = (s, s) + (s^\perp, s^\perp),$$

siendo nulos los restantes términos ya que s y $s^\perp$ son ortogonales. Esto demuestra (1.18).

DEFINICIÓN. *Sea S un subespacio de dimensión finita de un espacio euclídeo V, y sea* $\{e_1, \ldots, e_n\}$ *una base ortonormal para S. Si* $x \in V$, *el elemento s definido por la ecuación*

$$s = \sum_{i=1}^{n} (x, e_i)e_i$$

se denomina proyección de x sobre el subespacio S.

Demostramos seguidamente que la proyección de x sobre S es la solución del problema de aproximación establecido al comienzo de esta sección.

1.16 Aproximación óptima de elementos de un espacio euclídeo por elementos de un subespacio de dimensión finita

TEOREMA 1.16. TEOREMA DE APROXIMACIÓN. *Sea S un subespacio de dimensión finita de un espacio euclídeo V, y sea x un elemento de V. La proyección de x sobre S es más próxima a x que cualquier otro elemento de S. Esto es, si s es la proyección de x sobre S, tenemos*

$$\|x - s\| \le \|x - t\|$$

para todo t de S; es válido el signo de igualdad si y sólo si $t = s$.

Demostración. En virtud del teorema 1.15 podemos escribir $x = s + s^\perp$, donde $s \in S$ y $s^\perp \in S^\perp$. Entonces, para cualquier t de S, tenemos

$$x - t = (x - s) + (s - t).$$

Puesto que $s - t \in S$ y $x - s = s^\perp \in S^\perp$, ésta es una descomposición ortogonal de $x - t$, así que su norma viene dada por la fórmula pitagórica

$$\|x - t\|^2 = \|x - s\|^2 + \|s - t\|^2.$$

Pero $\|s - t\|^2 \geq 0$, con lo que $\|x - t\|^2 \geq \|x - s\|^2$, valiendo el signo igual si y sólo si $s = t$. Esto completa la demostración.

EJEMPLO 1. *Aproximación de funciones continuas en* $[0, 2\pi]$, *por polinomios trigonométricos.* Sea $V = C(0, 2\pi)$, el espacio lineal de todas las funciones reales continuas en el intervalo $[0, 2\pi]$, y definamos un producto interior mediante la ecuación $(f, g) = \int_0^{2\pi} f(x)g(x)\,dx$. En la sección 1.12 vimos un conjunto ortonormal de funciones trigonométricas $\varphi_0, \varphi_1, \varphi_2, \ldots,$ donde

$$(1.21) \qquad \varphi_0(x) = \frac{1}{\sqrt{2\pi}}, \qquad \varphi_{2k-1}(x) = \frac{\cos kx}{\sqrt{\pi}}, \qquad \varphi_{2k}(x) = \frac{\operatorname{sen} kx}{\sqrt{\pi}}, \qquad \text{para } k \geq 1\,.$$

Los $2n + 1$ elementos $\varphi_0, \varphi_1, \ldots, \varphi_{2n}$ generan un subespacio S de dimensión $2n + 1$. Los elementos de S se llaman *polinomios trigonométricos.*

Si $f \in C(0, 2\pi)$, sea f_n la proyección de f sobre el subespacio S. Tenemos entonces

$$(1.22) \qquad f_n = \sum_{k=0}^{2n} (f, \varphi_k)\varphi_k\,, \qquad \text{donde} \quad (f, \varphi_k) = \int_0^{2\pi} f(x)\varphi_k(x)\,dx\,.$$

Los números (f, φ_k) se llaman *coeficientes de Fourier* de f. Aplicando las fórmulas (1.21), podemos poner (1.22) en la forma

$$(1.23) \qquad f_n(x) = \tfrac{1}{2}a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \operatorname{sen} kx)\,,$$

donde

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,dx\,, \qquad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\operatorname{sen} kx\,dx$$

para $k = 0, 1, 2, \ldots, n$. El teorema de aproximación nos dice que el polinomio trigonométrico (1.23) aproxima f mejor que cualquier otro polinomio trigonométrico de S, en el sentido de que la norma $\|f - f_n\|$ es la más pequeña posible.

EJEMPLO 2. *Aproximación de funciones continuas en* $[-1, 1]$ *por polinomios de grado* $\leq n$. Sea $V = C(-1, 1)$, el espacio de las funciones reales continuas en $[-1, 1]$, y sea $(f, g) = \int_{-1}^{1} f(x)\,g(x)\,dx$. Los $n + 1$ polinomios de Legendre normalizados $\varphi_0, \varphi_1, \ldots, \varphi_n$, introducidos en la sección 1.14, generan un subespacio S de dimensión $n + 1$ que consta de todos los polinomios de grado

$\leq n$. Si $f \in C(-1, 1)$, designemos con f_n la proyección de f sobre S. Tenemos entonces

$$f_n = \sum_{k=0}^{n} (f, \varphi_k)\varphi_k , \qquad \text{donde} \quad (f, \varphi_k) = \int_{-1}^{1} f(t)\varphi_k(t)\, dt .$$

Este es el polinomio de grado $\leq n$ para el que la norma $\|f - f_n\|$ es la menor. Por ejemplo, cuando $f(x) = \operatorname{sen} \pi x$, los coeficientes (f, φ_k) vienen dados por

$$(f, \varphi_k) = \int_{-1}^{1} \operatorname{sen} \pi t \; \varphi_k(t)\, dt .$$

En particular, tenemos $(f, \varphi_0) = 0$ y

$$(f, \varphi_1) = \int_{-1}^{1} \sqrt{\frac{3}{2}}\, t \operatorname{sen} \pi t \, dt = \sqrt{\frac{3}{2}}\, \frac{2}{\pi} .$$

Por consiguiente el polinomio de primer grado $f_1(t)$ más próximo a $\operatorname{sen} \pi t$ en $[-1, 1]$ es

$$f_1(t) = \sqrt{\frac{3}{2}}\, \frac{2}{\pi}\, \varphi_1(t) = \frac{3}{\pi}\, t .$$

Puesto que $(f, \varphi_2) = 0$, este es también la mejor aproximación cuadrática.

1.17 Ejercicios

1. En cada caso, hallar una base ortonormal para el subespacio de V_3 generado por los vectores dados.
   a) $x_1 = (1, 1, 1), \quad x_2 = (1, 0, 1), \quad x_3 = (3, 2, 3)$.
   b) $x_1 = (1, 1, 1), \quad x_2 = (-1, 1, -1), \quad x_3 = (1, 0, 1)$.
2. En cada caso, hallar una base ortonormal para el subespacio de V_4 generado por los vectores dados.
   a) $x_1 = (1\ \ 1, 0, 0), \quad x_2 = (0, 1, 1, 0), \quad x_3 = (0, 0, 1, 1), \quad x_4 = (1, 0, 0, 1)$.
   b) $x_1 = (1, 1, 0, 1), \quad x_2 = (1, 0, 2, 1), \quad x_3 = (1, 2, -2, 1)$.
3. En el espacio lineal real $C(0, \pi)$ con producto interior $(x, y) = \int_0^{\pi} x(t)y(t)dt$, sea $x_n(t) = \cos nt$ para $n = 0, 1, 2, \ldots$. Demostrar que las funciones $y_0, y_1, y_2, \ldots$, dadas por

$$y_0(t) = \frac{1}{\sqrt{\pi}} \qquad \text{y} \qquad y_n(t) = \sqrt{\frac{2}{\pi}} \cos nt \qquad \text{para} \quad n \geq 1 ,$$

forman un conjunto ortonormal que genera el mismo subespacio que $x_0, x_1, x_2, \ldots$

4. En el espacio lineal de todos los polinomios reales, con producto interior $(x, y) = \int_0^1 x(t)y(t)dt$, sea $x_n(t) = t^n$ para $n = 0, 1, 2, \ldots$. Demostrar que las funciones

$$y_0(t) = 1\,, \qquad y_1(t) = \sqrt{3}\,(2t - 1)\,, \qquad y_2(t) = \sqrt{5}\,(6t^2 - 6t + 1)$$

forman un conjunto ortonormal que genera el mismo subespacio que $\{x_0, x_1, x_2\}$.
5. Sea V el espacio lineal de todas las funciones reales f continuas en $[0, +\infty)$ y tales que la integral $\int_0^\infty e^{-t}f^2(t)dt$ converge. Definamos $(f, g) = \int_0^\infty e^{-t}f(t)g(t)dt$, y sea $y_0, y_1, y_2, \ldots$, el conjunto obtenido aplicando el método de Gram-Schmidt a $x_0, x_1, x_2, \ldots$, donde $x_n(t) = t^n$ para $n \geq 0$. Demostrar que $y_0(t) = 1$, $y_1(t) = t - 1$, $y_2(t) = t^2 - 4t + 2$, $y_3(t) = t^3 - 9t^2 + 18t - 6$.
6. En el espacio lineal real $C(1, 3)$ con producto interior $(f, g) = \int_1^3 f(x)g(x)dx$, sea $f(x) = 1/x$ y demostrar que el polinomio constante g más próximo a f es $g = \frac{1}{2}\log 3$. Calcular $\|g - f\|^2$ para este g.
7. En el espacio lineal real $C(0, 2)$ con producto interior $(f, g) = \int_0^2 f(x)g(x)dx$, sea $f(x) = e^x$ y demostrar que el polinomio constante g más próximo a f es $g = \frac{1}{2}(e^2 - 1)$. Calcular $\|g - f\|^2$ para este g.
8. En el espacio lineal real $C(-1, 1)$ con producto interior $(f, g) = \int_{-1}^1 f(x)g(x)dx$, sea $f(x) = e^x$ y hallar el polinomio g más próximo a f. Calcular $\|g - f\|^2$ para este g.
9. En el espacio lineal real $C(0, 2\pi)$ con producto interior $(f, g) = \int_0^{2\pi} f(x)g(x)dx$, sea $f(x) = x$. En el subespacio generado por $u_0(x) = 1$, $u_1(x) = \cos x$, $u_2(x) = \operatorname{sen} x$, hallar el polinomio trigonométrico más próximo a f.
10. En el espacio lineal V del ejercicio 5, poner $f(x) = e^{-x}$ y hallar el polinomio de primer grado más próximo a f.

2

TRANSFORMACIONES LINEALES Y MATRICES

2.1 Transformaciones lineales

Uno de los modernos objetivos del Análisis es un estudio amplio de funciones cuyos dominios y recorridos son subconjuntos de espacios lineales. Tales funciones se llaman *transformaciones, aplicaciones,* u *operadores.* Este capítulo trata de los ejemplos más sencillos, llamados transformaciones *lineales,* que se presentan en todas las ramas de la Matemática. Las propiedades de transformaciones más generales se obtienen a menudo aproximándolas mediante transformaciones lineales.

Introducimos primero la notación y la terminología más corriente relativa a funciones cualesquiera. Sean V y W dos conjuntos. El símbolo

$$T : V \to W$$

se usará para indicar que T es una función cuyo dominio es V y cuyos valores están en W. Para cada x de V, el elemento $T(x)$ de W se llama *imagen de x a través de T,* y decimos que *T aplica x en $T(x)$*. Si A es un subconjunto cualquiera de V, el conjunto de todas las imágenes $T(x)$ para x de A se llama *la imagen de A a través de T* y se representa por $T(A)$. La imagen del dominio V, $T(V)$, es el recorrido de T.

Supongamos ahora que V y W son espacios lineales que tienen el mismo conjunto de escalares, y definamos una transformación lineal como sigue.

DEFINICIÓN. *Si V y W son dos espacios lineales, una función $T: V \to W$ se llama transformación lineal de V en W, si tiene las propiedades siguientes:*

a) $T(x + y) = T(x) + T(y)$ *cualesquiera que sean x e y de V,*
b) $T(cx) = cT(x)$ *para todo x de V y cualquier escalar c.*

Esto significa que T conserva la adición y la multiplicación por escalares. Las dos propiedades pueden combinarse en una fórmula que establece que

$$T(ax + by) = aT(x) + bT(y)$$

para todo x y todo y de V y todos los escalares a y b. Por inducción, tenemos también la relación más general

$$T\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i T(x_i)$$

para n elementos cualesquiera $x_1, \ldots, x_n$ de V, y n escalares cualesquiera $a_1, \ldots, a_n$.

El lector puede comprobar fácilmente que los ejemplos siguientes son transformaciones lineales.

EJEMPLO 1. *Transformación idéntica.* La transformación $T: V \to V$, donde $T(x) = x$ para todo x de V, se denomina transformación idéntica y se designa por I o por I_V.

EJEMPLO 2. *Transformación cero.* La transformación $T: V \to V$ que aplica cada elemento de V en O se llama transformación cero y se designa por O.

EJEMPLO 3. *Multiplicación por un escalar fijo c.* Tenemos aquí $T: V \to V$, donde $T(x) = cx$ para todo x de V. Cuando $c = 1$, se trata de la transformación idéntica. Cuando $c = 0$, es la transformación cero.

EJEMPLO 4. *Ecuaciones lineales.* Sean $V = V_n$ y $W = V_m$. Dados mn números reales a_{ik}, con $i = 1, 2, \ldots, m$ y $k = 1, 2, \ldots, n$, definamos $T: V_n \to V_m$ como sigue: T aplica cada vector $x = (x_1, \ldots, x_n)$ de V_n en el vector $y = (y_1, \ldots, y_m)$ de V_m de acuerdo con las ecuaciones

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad \text{para} \quad i = 1, 2, \ldots, m .$$

EJEMPLO 5. *Producto interior con un elemento fijo.* Sea V un espacio real euclídeo. Para un elemento fijo z de V, definamos $T: V \to \mathbf{R}$ así: Si $x \in V$, $T(x) = (x, z)$, el producto interior de x por z.

EJEMPLO 6. *Proyección sobre un subespacio.* Sean V un espacio euclídeo y S un subespacio de V de dimensión finita. Definamos $T: V \to S$ así: Si $x \in V$, $T(x)$ es la proyección de x sobre S.

EJEMPLO 7. *El operador derivación.* Sea V el espacio lineal de todas las funciones reales f derivables en un intervalo abierto (a, b). La transformación lineal que aplica cada función f de V en su derivada f' se llama operador derivación y se designa por D. Así pues, tenemos $D: V \to W$, donde $D(f) = f'$ para cada f de V. El espacio W contiene todas las derivadas f'.

EJEMPLO 8. *El operador integracion.* Sea V el espacio lineal de todas las funciones reales continuas en un intervalo $[a, b]$. Si $f \in V$, definamos $g = T(f)$ como la función V dada por

$$g(x) = \int_a^x f(t)\,dt \qquad \text{si} \quad a \leq x \leq b\,.$$

Esta transformación T se llama operador integración.

2.2. Núcleo y recorrido

En esta sección, T representa una transformación lineal de un espacio lineal V en un espacio lineal W.

TEOREMA 2.1. *El conjunto $T(V)$ (recorrido de T) es un subespacio de W. Además, T aplica el elemento cero de V en el elemento cero de W.*

Demostración. Para demostrar que $T(V)$ es un subespacio de W, tan sólo necesitamos comprobar los axiomas de clausura. Tomemos dos elementos cualesquiera de $T(V)$, sean $T(x)$ y $T(y)$. Entonces $T(x) + T(y) = T(x + y)$, así que $T(x) + T(y)$ pertenece a $T(V)$. Asimismo, para cualquier escalar c tenemos $cT(x) = T(cx)$, con lo que $cT(x)$ pertenece a $T(V)$. Por consiguiente, $T(V)$ es un subespacio de W. Tomando $c = 0$ en la relación $T(cx) = cT(x)$, encontramos que $T(O) = O$.

DEFINICIÓN. *El conjunto de todos los elementos de V que T aplica en O se llama núcleo de T y se designa por $N(T)$. Así pues, tenemos*

$$N(T) = \{x \mid x \in V \quad y \quad T(x) = O\}\,.$$

TEOREMA 2.2. *El núcleo de T es un subespacio de V.*

Demostración. Si x e y están en $N(T)$, lo mismo les ocurre a $x + y$ y a cx para todos los escalares c, ya que

$$T(x + y) = T(x) + T(y) = O \qquad \text{y} \qquad T(cx) = cT(x) = O.$$

Los ejemplos que siguen describen los núcleos de las transformaciones lineales dadas en la sección 2.1.

EJEMPLO 1. *Transformación idéntica.* El núcleo es $\{O\}$, el subespacio constituido tan sólo por el elemento cero.

EJEMPLO 2. *Transformación cero.* Puesto que todo elemento de V se aplica en cero, el núcleo es el mismo V.

EJEMPLO 3. *Multiplicación por un escalar fijo c.* Si $c \neq 0$, el núcleo sólo contiene el elemento O. Si $c = 0$, el núcleo es V.

EJEMPLO 4. *Ecuaciones lineales.* El núcleo está constituido por todos los vectores $(x_1, \ldots, x_n)$ de V_n para los cuales

$$\sum_{k=1}^{n} a_{ik} x_k = 0 \quad \text{para} \quad i = 1, 2, \ldots, m .$$

EJEMPLO 5. *Producto interior por un elemento fijo z.* El núcleo consta de todos los elementos de V ortogonales a z.

EJEMPLO 6. *Proyección sobre un subespacio S.* Si $x \in V$, tenemos la única descomposición ortogonal $x = s + s^{\perp}$ (según el teorema 1.15). Puesto que $T(x) = s$, tenemos $T(x) = O$ si y sólo si $x = s^{\perp}$. Por consiguiente, el núcleo es $S^{\perp}$, el complemento ortogonal de S.

EJEMPLO 7. *Operador derivación.* El núcleo está formado por todas las funciones constantes en el intervalo dado.

EJEMPLO 8. *Operador integración.* El núcleo contiene solamente la función cero.

2.3 Dimensión del núcleo y rango de la transformación

También en esta sección T representa una transformación de un espacio lineal V en un espacio lineal W. Nos interesa la relación entre las dimensiones de V, del núcleo $N(T)$ y del recorrido $T(V)$. Si V es de dimensión finita, el núcleo también lo será por ser un subespacio de V. En el teorema que sigue, demostramos que el recorrido $T(V)$ también es de dimensión finita; su dimensión se llama *rango* de T.

TEOREMA 2.3. *Si V es de dimensión finita, también lo es $T(V)$, y tenemos*

$$\dim N(T) + \dim T(V) = \dim V . \tag{2.1}$$

Dicho de otro modo, la dimensión del núcleo más el rango de una transformación lineal es igual a la dimensión de su dominio.

Demostración. Sean $n = \dim V$ y $e_1, \ldots, e_k$ una base para $N(T)$, donde $k = \dim N(T) \leq n$. Según el teorema 1.7; esos elementos forman parte de una cierta base de V, por ejemplo de la base

$$e_1, \ldots, e_k, e_{k+1}, \ldots, e_{k+r}, \tag{2.2}$$

donde $k + r = n$. Demostraremos que los r elementos

$$T(e_{k+1}), \ldots, T(e_{k+r}) \tag{2.3}$$

forman una base de $T(V)$, demostrando así que $\dim T(V) = r$. Puesto que $k+r=n$, también eso demuestra (2.1).

Demostramos primero que los r elementos de (2.3) generan $T(V)$. Si $y \in T(V)$, es $y = T(x)$ para un cierto x de V, y podemos escribir $x = c_1e_1 + \ldots + c_{k+r}e_{k+r}$. Luego, tenemos

$$y = T(x) = \sum_{i=1}^{k+r} c_iT(e_i) = \sum_{i=1}^{k} c_iT(e_i) + \sum_{i=k+1}^{k+r} c_iT(e_i) = \sum_{i=k+1}^{k+r} c_iT(e_i)$$

puesto que $T(e_1) = \ldots = T(e_k) = O$. Esto demuestra que los elementos (2.3) generan $T(V)$.

Demostremos ahora que esos elementos son independientes. Supongamos que existieran escalares $c_{k+1}, \ldots, c_{k+r}$ tales que

$$\sum_{i=k+1}^{k+r} c_iT(e_i) = O .$$

Esto implicaría que

$$T\left(\sum_{i=k+1}^{k+r} c_ie_i\right) = O$$

por lo que el elemento $x = c_{k+1}e_{k+1} + \ldots + c_{k+r}e_{k+r}$ sería del núcleo $N(T)$. Significa esto que existirían escalares $c_1, \ldots, c_k$ tales que $x = c_1e_1 + \ldots + c_ke_k$, con lo que tendríamos

$$x - x = \sum_{i=1}^{k} c_ie_i - \sum_{i=k+1}^{k+r} c_ie_i = O .$$

Pero como los elementos (2.2) son independientes, todos los escalares c_i han de ser cero. Por consiguiente, los elementos (2.3) son independientes.

Nota: Si V es de dimensión infinita por lo menos uno de los dos $N(T)$ o $T(V)$ es de dimensión infinita. En el ejercicio 30 de la Sección 2.4 se esboza una demostración de este hecho.

2.4 Ejercicios

En cada uno de los ejercicios del 1 al 10, se define una transformación $T: V_2 \to V_2$ mediante la fórmula dada para $T(x, y)$, donde (x, y) es un punto cualquiera de V_2. Determinar en cada caso si T es lineal. Si T es lineal, decir cuáles son el núcleo y el recorrido, y calcular sus dimensiones

1. $T(x, y) = (y, x)$.
2. $T(x, y) = (x, -y)$.
3. $T(x, y) = (x, 0)$.
4. $T(x, y) = (x, x)$.
5. $T(x, y) = (x^2, y^2)$.
6. $T(x, y) = (e^x, e^y)$.
7. $T(x, y) = (x, 1)$.
8. $T(x, y) = (x + 1, y + 1)$.
9. $T(x, y) = (x - y, x + y)$.
10. $T(x, y) = (2x - y, x + y)$.

Hacer lo mismo en cada uno de los ejercicios del 11 al 15. Si la transformación $T: V_2 \to V_2$ es la que se indica.

11. T hace girar cualquier punto el mismo ángulo ϕ alrededor del origen. Esto es, T aplica un punto de coordenadas polares (r, θ) en el punto de coordenadas polares $(r, \theta + \phi)$, donde ϕ es fijo. Además, T aplica O en sí mismo.
12. T aplica cada punto en su simétrico respecto a una recta fija que pasa por el origen.
13. T aplica todo punto en el punto (1, 1).
14. T aplica cada punto de coordenadas polares (r, θ) en el punto de coordenadas $(2r, \theta)$. Además, T aplica O en sí mismo.
15. T aplica cada punto de coordenadas polares (r, θ) en el punto de coordenadas $(r, 2\theta)$. Además, T aplica O en sí mismo.

Hacer lo mismo en cada uno de los ejercicios 16 al 23 si la transformación $T: V_3 \to V_3$ está definida por la fórmula que se da para $T(x, y, z)$, donde (x, y, z) es un punto arbitrario de V_3.

16. $T(x, y, z) = (z, y, x)$.
17. $T(x, y, z) = (x, y, 0)$.
18. $T(x, y, z) = (x, 2y, 3z)$.
19. $T(x, y, z) = (x, y, 1)$.
20. $T(x, y, z) = (x + 1, y + 1, z - 1)$.
21. $T(x, y, z) = (x + 1, y + 2, z + 3)$.
22. $T(x, y, z) = (x, y^2, z^3)$.
23. $T(x, y, z) = (x + z, 0, x + y)$.

En cada uno de los ejercicios del 24 al 27, la transformación $T: V \to V$ es la que se indica. Determinar, en cada caso, si T es lineal. Si lo es, decir cuáles son el núcleo y el recorrido y calcular sus dimensiones cuando sean finitas.

24. Sea V el espacio lineal de todos los polinomios reales $p(x)$ de grado $\leq n$. Si $p \in V$, $q = T(p)$ significa que $q(x) = p(x+1)$ para todo real x.
25. Sea V el espacio lineal de todas las funciones reales derivables en el intervalo abierto $(-1, 1)$. Si $f \in V$, $g = T(f)$ significa que $g(x) = xf'(x)$ para todo x de $(-1, 1)$.
26. Sea V el espacio lineal de todas las funciones reales continuas en $[a, b]$. Si $f \in V$, $g = T(f)$ significa que

$$g(x) = \int_a^b f(t) \operatorname{sen}(x - t)\, dt \qquad \text{para} \quad a \leq x \leq b\,.$$

27. Sea V el espacio de todas las funciones reales derivables dos veces en un intervalo abierto (a, b). Si $y \in V$, definir $T(y) = y'' + Py' + Qy$ siendo P y Q dos constantes.
28. Sea V el espacio lineal de todas las sucesiones reales convergentes $\{x_n\}$. Definimos una transformación $T: V \to V$ así: Si $x = \{x_n\}$ es una sucesión convergente con límite a, ponemos $T(x) = \{y_n\}$, donde $y_n = a - x_n$ para $n \geq 1$. Demostrar que T es lineal y decir cuáles son el núcleo y el recorrido de T.
29. Sea V el espacio lineal de todas las funciones continuas en el intervalo $[-\pi, \pi]$. Sea S el subconjunto de V que consta de todas las funciones f que satisfacen las tres ecuaciones

$$\int_{-\pi}^{\pi} f(t)\, dt = 0\,, \qquad \int_{-\pi}^{\pi} f(t) \cos t\, dt = 0\,, \qquad \int_{-\pi}^{\pi} f(t) \operatorname{sen} t\, dt = 0\,.$$

a) Demostrar que S es un subespacio de V.
b) Demostrar que S contiene las funciones $f(x) = \cos nx$ y $f(x) = \operatorname{sen} nx$ para cada $n = 2, 3, \ldots$
c) Demostrar que S es de dimensión infinita.

Sea $T: V \to V$ la transformación lineal definida así: Si $f \in V$, $g = T(f)$ significa que

$$g(x) = \int_{-\pi}^{\pi} \{1 + \cos(x - t)\} f(t)\, dt\,.$$

d) Demostrar que $T(V)$, el recorrido de T, es de dimensión finita y hallar una base para $T(V)$.
e) Determinar el núcleo de T.
f) Hallar todos los números reales $c \neq 0$ y todas las funciones f no nulas de V tales que $T(f) = cf$. (Obsérvese que una tal f pertenece al recorrido de T.)

30. Sea $T: V \to W$ una transformación lineal de un espacio lineal V en un espacio lineal W. Si V es de dimensión infinita, demostrar que por lo menos uno de los dos $T(V)$ o $N(T)$ es de dimensión infinita.

[*Indicación:* Supóngase $\dim N(T) = k$, $\dim T(V) = r$, sea $e_1, \ldots, e_k$ una base para $N(T)$ y sean $e_1, \ldots, e_k, e_{k+1}, \ldots e_{k+n}$ elementos independientes de V, siendo $n > r$. Los elementos $T(e_{k+1}), \ldots, T(e_{k+n})$ son dependientes ya que $n > r$. Utilizar este hecho para obtener una contradicción.]

2.5 Operaciones algebraicas con transformaciones lineales

Las funciones cuyos valores pertenecen a un espacio lineal dado W pueden sumarse unas con otras y pueden multiplicarse por escalares de W de acuerdo con la definición siguiente.

DEFINICIÓN. *Sean $S: V \to W$ y $T: V \to W$ dos funciones con un dominio común V y con valores pertenecientes a un espacio lineal W. Si c es un escalar cualquiera de W, definimos la suma $S + T$ y el producto cT por las ecuaciones*

(2.4) $$(S + T)(x) = S(x) + T(x), \qquad (cT)(x) = cT(x)$$

para todo x de V.

Nos interesa especialmente el caso en el que V es también un espacio lineal con los mismos escalares que W. En este caso designamos con $\mathscr{L}(V, W)$ el conjunto de todas las transformaciones lineales de V en W.

Si S y T son dos transformaciones lineales de $\mathscr{L}(V, W)$, es un sencillo ejercicio comprobar que $S+T$ y cT también son transformaciones lineales de $\mathscr{L}(V, W)$. Aún más. Con las operaciones que acabamos de definir, el mismo conjunto $\mathscr{L}(V, W)$ se transforma en un nuevo espacio lineal. La transformación cero sirve de elemento cero en ese espacio, y la transformación $(-1)T$ es la opuesta de T. Se comprueba que se satisfacen los diez axiomas de un espacio lineal. Por consiguiente, tenemos el siguiente.

TEOREMA 2.4. *El conjunto $\mathscr{L}(V, W)$ de todas las transformaciones lineales de V en W es un espacio lineal con las operaciones de adición y multiplicación por escalares definidas en* (2.4).

Una operación algebraica más interesante que se efectúa con las transformaciones lineales es la *composición o multiplicación* de transformaciones. Esta operación no utiliza la estructura algebraica de un espacio lineal y puede definirse con entera generalidad del siguiente modo.

DEFINICIÓN. *Dados los conjuntos U, V, W. Sean $T: U \to V$ una función con dominio U y valores en V, y $S: V \to W$ otra función con dominio V y valores en W. La composición ST es la función $ST: U \to W$ definida por*

$$(ST)(x) = S[T(x)] \quad \textit{para todo } x \textit{ en } U.$$

Así pues, para aplicar x mediante la composición ST, aplicamos primero x mediante T y luego aplicamos $T(x)$ por medio de S. Esto se representa en la figura 2.1.

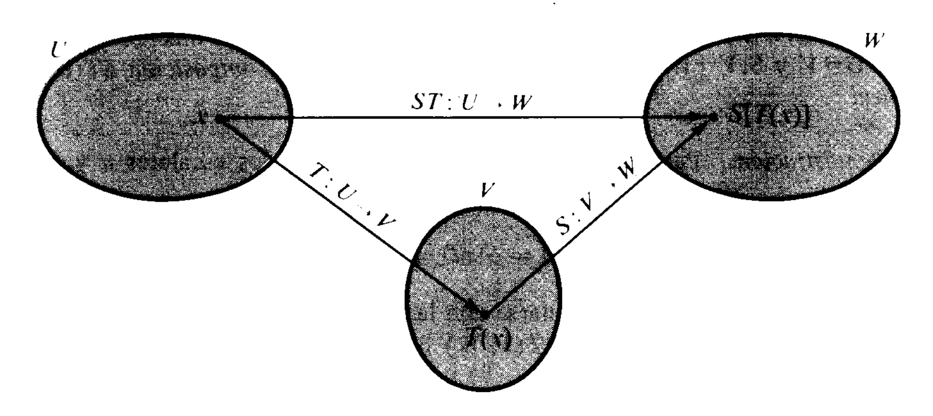


FIGURA 2.1 *Gráfico de la composición de dos transformaciones.*

La composición de funciones reales se ha encontrado repetidas veces en nuestro estudio del Cálculo, y hemos visto que la operación, en general, no es conmutativa. No obstante, como en el caso de las funciones reales, la composición satisface la ley asociativa.

TEOREMA 2.5. *Si $T: U \to V$, $S: V \to W$, y $R: W \to X$ son tres funciones, tenemos*

$$R(ST) = (RS)T.$$

Demostración. Las funciones $R(ST)$ y $(RS)T$ tienen ambas dominio U y valores en X. Para cada x de U, tenemos

$$[R(ST)](x) = R[(ST)(x)] = R[S[T(x)]] \quad \text{y} \quad [(RS)T](x) = (RS)[T(x)] = R[S[T(x)]],$$

lo que demuestra que $R(ST) = (RS)T$.

DEFINICIÓN. Sea $T: V \to V$ *una función que aplica V en sí mismo. Definimos inductivamente las potencias enteras de T como sigue:*

$$T^0 = I, \qquad T^n = TT^{n-1} \quad \text{para } n \geq 1.$$

Aquí I representa la transformación idéntica. El lector puede comprobar que la ley asociativa implica la ley de exponentes $T^mT^n = T^{m+n}$ para todos los enteros no negativos m y n.

El teorema que sigue prueba que la composición de transformaciones *lineales* es lineal.

TEOREMA 2.6. *Si U, V, W son espacios lineales con los mismos escalares, y si $T: U \to V$ y $S: V \to W$ son transformaciones lineales, la composición $ST: U \to W$ es lineal.*

Demostración. Para todo x y todo y de U y todos los escalares a y b, tenemos

$$(ST)(ax + by) = S[T(ax + by)] = S[aT(x) + bT(y)] = aST(x) + bST(y) .$$

La composición puede combinarse con las operaciones algebraicas de adición y multiplicación por escalares en $\mathscr{L}(V, W)$ llegando al siguiente

TEOREMA 2.7. *Sean U, V, W espacios lineales con los mismos escalares, supongamos que S y T pertenecen a $\mathscr{L}(V, W)$, y sea c un escalar cualquiera.*

a) *Para cualquier función R con valores en V, tenemos*

$$(S + T)R = SR + TR \qquad y \qquad (cS)R = c(SR) .$$

b) *Para cualquier transformación lineal $R: W \to U$, tenemos*

$$R(S + T) = RS + RT \qquad y \qquad R(cS) = c(RS) .$$

La demostración es una consecuencia inmediata de la definición de composición y se deja como ejercicio.

2.6 Inversas

Al estudiar las funciones reales aprendimos cómo construir nuevas funciones mediante la inversión de funciones monótonas. Queremos ahora extender el método de inversión a una clase más general de funciones.

Dada una función T, nuestro objetivo es encontrar, si es posible, otra función S cuya composición con T sea la transformación idéntica. Puesto que la composición, en general, no es conmutativa, tenemos que distinguir ST de TS. Por lo tanto introducimos dos tipos de inversas que llamamos inversa por la derecha e inversa por la izquierda.

DEFINICIÓN. *Dados dos conjuntos V y W y una función $T: V \to W$. Se dice que una función $S: T(V) \to V$ es inversa de T por la izquierda si $S[T(x)] = x$ para todo x de V, esto es, si*

$$ST = I_V ,$$

donde I_V es la transformación idéntica sobre V. Una función $R: T(V) \to V$ se llama inversa de T por la derecha si $T[R(y)] = y$ para todo y de $T(V)$, esto es, si

$$TR = I_{T(V)},$$

donde $I_{T(V)}$ es la transformación idéntica sobre $T(V)$.

EJEMPLO. *Una función sin inversa por la izquierda pero con dos inversas por la derecha.* Sean $V = \{1, 2\}$ y $W = \{0\}$. Definimos $T: V \to W$ como sigue: $T(1) = T(2) = 0$. Esta función tiene dos inversas por la derecha $R: W \to V$ y $R': W \to V$ dadas por

$$R(0) = 1, \qquad R'(0) = 2.$$

No puede tener inversa por la izquierda S ya que ello exigiría

$$1 = S[T(1)] = S(0) \qquad \text{y} \qquad 2 = S[T(2)] = S(0).$$

Este sencillo ejemplo pone de manifiesto que no tiene que existir necesariamente inversa por la izquierda y que la inversa por la derecha no tiene que ser necesariamente única.

Toda función $T: V \to W$ tiene por lo menos una inversa a la derecha. En efecto, cada y de $T(V)$ tiene la forma $y = T(x)$ para al menos un x de V. Si elegimos uno de esos valores x y definimos $R(y) = x$, entonces $T[R(y)] = T(x) = y$ para cada y de $T(V)$, así que R es una inversa por la derecha. La no unicidad puede presentarse debido a que puede haber más de un x de V que se aplique en un y de $T(V)$. Dentro de poco demostraremos (teorema 2.9) que si cada y de $T(V)$ es la imagen de *un sólo* x de V, la inversa por la derecha es única.
Antes demostraremos que si existe inversa por la izquierda es única y, al mismo tiempo, es inversa a la derecha.

TEOREMA 2.8. *Una $T: V \to W$ puede tener a lo más una inversa por la izquierda. Si T tiene inversa por la izquierda S, entonces S es también inversa por la derecha.*

Demostración. Supongamos que T tenga dos inversas por la izquierda, $S: T(V) \to V$ y $S': T(V) \to V$. Elijamos cualquier y en $T(V)$. Demostraremos que $S(y) = S'(y)$. Como $y = T(x)$ para un cierto x de V, tenemos

$$S[T(x)] = x \qquad \text{y} \qquad S'[T(x)] = x,$$

puesto que S y S' son ambas inversas por la izquierda. Por consiguiente $S(y) = x$ y $S'(y) = x$, con lo que $S(y) = S'(y)$ para todo y de $T(V)$. Por lo tanto $S = S'$ lo que demuestra que las inversas por la izquierda coinciden.

Demostremos ahora que toda inversa por la izquierda S es también inversa por la derecha. Elijamos un elemento cualquiera y en $T(V)$. Demostraremos que $T[S(y)] = y$. Puesto que $y \in T(V)$, tenemos $y = T(x)$ para un cierto x de V. Pero S es inversa por la izquierda, así que

$$x = S[T(x)] = S(y).$$

Aplicando T, llegamos a $T(x) = T[S(y)]$. Pero $y = T(x)$, con lo que $y = T[S(y)]$, lo cual completa la demostración.

El teorema que sigue caracteriza todas las funciones que tienen inversa por la izquierda.

TEOREMA 2.9. *Una función $T: V \to W$ tiene inversa por la izquierda si y sólo si T aplica elementos distintos de V en elementos distintos de W; esto es, si y sólo si, para cualesquiera x e y de V,*

$$x \neq y \quad \textit{implica} \quad T(x) \neq T(y). \tag{2.5}$$

Nota: La condición (2.5) es equivalente a la afirmación

$$T(x) = T(y) \quad \text{implica} \quad x = y. \tag{2.6}$$

Una función T que satisface (2.5) o (2.6) para cualesquiera x e y de V se denomina *uno a uno* en V.

Demostración. Supongamos que S es la inversa por la izquierda de T, y que $T(x) = T(y)$. Queremos demostrar que $x = y$. Aplicando S, encontramos $S[T(x)] = S[T(y)]$. Puesto que $S[T(x)] = x$ y $S[T(y)] = y$, esto implica $x = y$. Con ello queda demostrado que una función con inversa por la izquierda es uno a uno en su dominio.

Demostremos ahora el recíproco. Supongamos que T es uno a uno en V. Encontraremos una función $S: T(V) \to V$ que es inversa de T por la izquierda. Si $y \in T(V)$, entonces $y = T(x)$ para un cierto x de V. En virtud de (2.6), existe *exactamente un* x en V para el cual $y = T(x)$. Definamos $S(y)$ como ese x. Esto es, definamos S en $T(V)$ como sigue:

$$S(y) = x \quad \text{implica que} \quad T(x) = y.$$

Tenemos entonces $S[T(x)] = x$ para cada x de V, así que $ST = I_V$. Por consiguiente, la función S así definida es inversa de T por la izquierda.

DEFINICIÓN. *Sea $T: V \to W$ uno a uno en V. La única inversa de T por la izquierda (la cual sabemos que también es inversa por la derecha) se designa por T^{-1}. Decimos que T es invertible, y llamamos a T^{-1} la inversa de T.*

Los resultados de esta sección se refieren a funciones cualesquiera. Seguidamente aplicamos esas ideas a las transformaciones lineales.

2.7 Transformaciones lineales uno a uno

En esta sección, V y W representan espacios lineales con los mismos escalares, y $T: V \to W$ es una transformación lineal de $\mathscr{L}(V, W)$. La linealidad de T nos permite expresar de varias maneras la propiedad de que una transformación lineal sea uno a uno.

TEOREMA 2.10. *Sea $T: V \to W$ una transformación lineal de $\mathscr{L}(V, W)$. Son equivalentes las siguientes proposiciones.*

a) *T es uno a uno en V.*

b) *T es invertible y su inversa $T^{-1}: T(V) \to V$ es lineal.*

c) *Para todo x de V, $T(x) = O$ implica $x = O$. Esto es, el núcleo $N(T)$ contiene solamente el elemento cero de V.*

Demostración. Demostraremos que a) implica b), b) implica c), y c) implica a). Supongamos primero que a) es cierta. T tiene entonces inversa (según el teorema 2.9), y tenemos que demostrar que T^{-1} es lineal. Tomemos dos elementos cualesquiera u y v de $T(V)$. Entonces $u = T(x)$ y $v = T(y)$ para algún x y algún y de V. Para dos escalares cualesquiera a y b, tenemos

$$au + bv = aT(x) + bT(y) = T(ax + by),$$

ya que T es lineal. Luego, aplicando T^{-1}, tenemos

$$T^{-1}(au + bv) = ax + by = aT^{-1}(u) + bT^{-1}(v),$$

así que T^{-1} es lineal. Por consiguiente a) implica b).

Supongamos seguidamente que b) es cierta. Tomemos un x cualquiera de V para el cual $T(x) = O$. Aplicando T^{-1}, encontramos que $x = T^{-1}(O) = O$, puesto que T^{-1} es lineal. Por consiguiente, b) implica c).

Por último, supongamos cierta c). Tomemos dos elementos cualesquiera u y v de V siendo $T(u) = T(v)$. Por la linealidad, tenemos $T(u - v) = T(u) - T(v) = O$, así que $u - v = O$. Por consiguiente, T es uno a uno en V, y queda completada la demostración del teorema.

Cuando V es de dimensión finita, la propiedad de ser uno a uno puede formularse en función de la dependencia y de la dimensionalidad, como se indica en el teorema que sigue.

TEOREMA 2.11. *Sea* $T: V \to W$ *una transformación lineal de* $\mathscr{L}(V, W)$ *y supongamos que* V *es de dimensión finita,* $\dim V = n$. *Entonces son equivalentes las proposiciones siguientes.*

a) T *es uno a uno en* V.

b) *Si* $e_1, \ldots, e_p$ *son elementos independientes de* V, $T(e_1), \ldots, T(e_p)$ *son elementos independientes de* $T(V)$.

c) $\dim T(V) = n$.

d) *Si* $\{e_1, \ldots, e_n\}$ *es una base para* V, $\{T(e_1), \ldots, T(e_n)\}$ *es una base para* $T(V)$.

Demostración. Probaremos que a) implica b), b) implica c), c) implica d), y d) implica a). Supongamos que a) es cierta. Sean $e_1, \ldots, e_p$ elementos independientes de V y consideremos los elementos $T(e_1), \ldots, T(e_p)$ de $T(V)$. Supongamos que

$$\sum_{i=1}^{p} c_i T(e_i) = O$$

para ciertos escalares $c_1, \ldots, c_p$. En virtud de la linealidad, obtenemos

$$T\left(\sum_{i=1}^{p} c_i e_i\right) = O\,, \quad \text{y por tanto} \quad \sum_{i=1}^{p} c_i e_i = O$$

ya que T es uno a uno. Pero $e_1, \ldots, e_p$ son independientes, así que $c_1 = \ldots = c_p = 0$. Por consiguiente a) implica b).

Supongamos ahora que es cierta b). Sea $\{e_1, \ldots, e_n\}$ una base para V. Según b), los n elementos $T(e_1), \ldots, T(e_n)$ de $T(V)$ son independientes. Por consiguiente, $\dim T(V) \geq n$. Pero, según el teorema 2.3, tenemos $\dim T(V) \leq n$. Luego $\dim T(V) = n$, con lo cual b) implica c).

Supongamos, seguidamente, que es cierta c) y sea $\{e_1, \ldots, e_n\}$ una base para V. Tomemos un elemento cualquiera y en $T(V)$. Entonces $y = T(x)$ para algún x en V, así que tenemos

$$x = \sum_{i=1}^{n} c_i e_i\,, \quad \text{y por tanto} \quad y = T(x) = \sum_{i=1}^{n} c_i T(e_i)\,.$$

Por consiguiente $\{T(e_1), \ldots, T(e_n)\}$ genera $T(V)$. Pero hemos supuesto que $\dim T(V) = n$, así que $\{T(e_1), \ldots, T(e_n)\}$ es una base para $T(V)$. Por consiguiente c) implica d).

Por último, supongamos cierta d). Demostraremos que $T(x) = O$ implica $x = O$. Sea $\{e_1, \ldots, e_n\}$ una base para V. Si $x \in V$, podemos escribir

$$x = \sum_{i=1}^{n} c_i e_i, \quad \text{y por tanto} \quad T(x) = \sum_{i=1}^{n} c_i T(e_i).$$

Si $T(x) = O$, entonces $c_1 = \ldots = c_n = 0$, puesto que los elementos $T(e_1), \ldots, T(e_n)$ son independientes. Por lo tanto $x = O$, con lo cual T es uno a uno en V. Así pues, d) implica a) y el teorema queda demostrado.

2.8 Ejercicios

1. Sea $V = \{0, 1\}$. Describir todas las funciones $T: V \to V$. En total son cuatro. Desígnense con T_1, T_2, T_3, T_4 y construir una tabla de multiplicación que muestre la composición de cada par. Indicar cuáles son uno a uno en V y dar sus inversas.
2. Sea $V = \{0, 1, 2\}$. Describir todas las funciones $T: V \to V$ para las cuales $T(V) = V$. En total son seis. Desígnense con $T_1, T_2, \ldots, T_6$ y construir una tabla de multiplicación que muestre la composición de cada par. Indicar cuáles son uno a uno en V, y dar sus inversas.

En cada uno de los ejercicios del 3 al 12, una función $T: V_2 \to V_2$ se define con la fórmula que se da para $T(x, y)$, siendo (x, y) un punto cualquiera de V_2. Determinar en cada caso si T es uno a uno en V_2. Si es así, describir su recorrido $T(V_2)$; para cada punto (u, v) de $T(V_2)$, poner $(x, y) = T^{-1}(u, v)$ y dar fórmulas para determinar x e y en función de u y v.

3. $T(x, y) = (y, x)$.
4. $T(x, y) = (x, -y)$.
5. $T(x, y) = (x, 0)$.
6. $T(x, y) = (x, x)$.
7. $T(x, y) = (x^2, y^2)$.
8. $T(x, y) = (e^x, e^y)$.
9. $T(x, y) = (x, 1)$.
10. $T(x, y) = (x + 1, y + 1)$.
11. $T(x, y) = (x - y, x + y)$.
12. $T(x, y) = (2x - y, x + y)$.

En cada uno de los ejercicios del 13 al 20, se define una función $T: V_3 \to V_3$ con la fórmula que se da para $T(x, y, z)$, siendo (x, y, z) un punto cualquiera de V_3. En cada caso, determinar si T es uno a uno en V_3. Si es así, describir su recorrido $T(V_3)$; para cada punto (u, v, w) de $T(V_3)$, póngase $(x, y, z) = T^{-1}(u, v, w)$ y dar fórmulas para la determinación de x, y, z en función de u, v, y w.

13. $T(x, y, z) = (z, y, x)$.
14. $T(x, y, z) = (x, y, 0)$.
15. $T(x, y, z) = (x, 2y, 3z)$.
16. $T(x, y, z) = (x, y, x + y + z)$.
17. $T(x, y, z) = (x + 1, y + 1, z - 1)$.
18. $T(x, y, z) = (x + 1, y + 2, z + 3)$.
19. $T(x, y, z) = (x, x + y, x + y + z)$.
20. $T(x, y, z) = (x + y, y + z, x + z)$.

21. Sea $T: V \to V$ una función que aplica V en sí mismo. Por inducción se definen las potencias con las fórmulas $T^0 = 1$, $T^n = TT^{n-1}$ para $n \geq 1$. Demostrar que la ley asociativa para la composición implica la ley de exponente: $T^m T^n = T^{m+n}$. Si T es invertible, demostrar que T^n también es invertible y que $(T^n)^{-1} = (T^{-1})^n$.

En los ejercicios del 22 al 25, S y T representan funciones con dominio V y valores en V. En general $ST \neq TS$. Si $ST = TS$, decimos que S y T *conmutan.*

22. Si S y T conmutan, demostrar que $(ST)^n = S^nT^n$ para cualquier entero $n \geq 0$.
23. Si S y T son invertibles, demostrar que ST también lo es y que $(ST)^{-1} = T^{-1}S^{-1}$. Dicho de otro modo, la inversa de ST es la composición de las inversas, tomadas en orden inverso.
24. Si S y T son invertibles y conmutan, demostrar que sus inversas también conmutan.
25. Sea V un espacio lineal. Si S y T conmutan, demostrar que

$$(S + T)^2 = S^2 + 2ST + T^2 \qquad \text{y} \qquad (S + T)^3 = S^3 + 3S^2T + 3ST^2 + T^3 .$$

Indicar cómo deben modificarse esas fórmulas si $ST \neq TS$.

26. Sean S y T las transformaciones lineales de V_3 en V_3 definidas por las fórmulas $S(x, y, z) = (z, y, x)$ y $T(x, y, z) = (x, x + y, x + y + z)$, siendo (x, y, z) un punto cualquiera de V_3
a) Determinar la imagen de (x, y, z) a través de cada una de las transformaciones siguientes: ST, TS, $ST - TS$, S^2, T^2, $(ST)^2$, $(TS)^2$, $(ST - TS)^2$.
b) Demostrar que S y T son uno a uno en V_3 y hallar la imagen de (u, v, w) a través de cada una de las transformaciones: S^{-1}, T^{-1}, $(ST)^{-1}$, $(TS)^{-1}$.
c) Hallar la imagen de (x, y, z) a través de $(T - I)^n$ para cada $n \geq 1$.
27. Sea V el espacio lineal de todos los polinomios reales $p(x)$. Sean D el operador derivación y T el operador integración que aplica cada polinomio p en un polinomio q dado por $q(x) = \int_0^x p(t)dt$. Demostrar que $DT = I_V$ pero que $TD \neq I_V$. Describir el núcleo y el recorrido de TD.
28. Sea V el espacio lineal de todos los polinomios reales $p(x)$. Sean D el operador derivación y T la transformación lineal que aplica $p(x)$ en $xp'(x)$.
a) Poner $p(x) = 2 + 3x - x^2 + 4x^3$ y determinar la imagen de p a través de cada una de las transformaciones siguientes: D, T, DT, TD, $DT - TD$, $T^2D^2 - D^2T^2$.
b) Determinar los polinomios p de V para los cuales $T(p) = p$.
c) Determinar los polinomios p de V para los cuales $(DT - 2D)(p) = O$.
d) Determinar los polinomios p de V para los cuales $(DT - TD)^n(p) = D^n(p)$.
29. Sean V y D como en el ejercicio 28 pero T es la transformación lineal que aplica $p(x)$ en $xp(x)$. Demostrar que $DT - TD = I$ y que $DT^n - T^nD = nT^{n-1}$ para todo $n \geq 2$.
30. Sean S y T dos transformaciones lineales de $\mathscr{L}(V, V)$ y supongamos que $ST - TS = I$. Demostrar que $ST^n - T^nS = nT^{n-1}$ para todo $n \geq 1$.
31. Sea V el espacio lineal de todos los polinomios $p(x)$. Sean R, S, T funciones que aplican un polinomio cualquiera $p(x) = c_0 + c_1x + \ldots + c_nx^n$ de V en los polinomios $r(x)$, $s(x)$ y $t(x)$ respectivamente, siendo

$$r(x) = p(0), \qquad s(x) = \sum_{k=1}^{n} c_kx^{k-1}, \qquad t(x) = \sum_{k=0}^{n} c_kx^{k+1}.$$

a) Poner $p(x) = 2 + 3x - x^2 + x^3$ y determinar la imagen de p a través de cada una de las transformaciones siguientes: R, S, T, ST, TS, $(TS)^2$, T^2S^2, S^2T^2, TRS, RST.
b) Demostrar que R, S y T son lineales y determinar el núcleo y el recorrido de cada una.
c) Demostrar que T es uno a uno en V y determinar su inversa.
d) Si $n \geq 1$, expresar $(TS)^n$ y S^nT^n en función de I y R.
32. En relación con el ejercicio 28 de la Sección 2.4. Determinar si T es uno a uno en V. Si lo es, decir cuál es la inversa.

2.9 Transformaciones lineales con valores asignados

Si V es de dimensión finita, siempre podemos construir una transformación lineal $T: V \to W$ con valores asignados a los elementos base de V, como se explica en el teorema siguiente.

TEOREMA 2.12. *Sea $e_1, \ldots, e_n$ una base para un espacio lineal n-dimensional V. Sean $u_1, \ldots, u_n$, n elementos arbitrarios de un espacio lineal W. Existe entonces una y sólo una transformación $T: V \to W$ tal que*

$$T(e_k) = u_k \quad \text{para} \quad k = 1, 2, \ldots, n. \tag{2.7}$$

Esta transformación T aplica un elemento cualquiera x de V del modo siguiente:

$$\text{Si} \quad x = \sum_{k=1}^{n} x_k e_k, \quad \text{entonces } T(x) = \sum_{k=1}^{n} x_k u_k. \tag{2.8}$$

Demostración. Todo x de V puede expresarse en forma única como combinación lineal de $e_1, \ldots, e_n$, siendo los multiplicadores $x_1, \ldots, x_n$ los componentes de x respecto a la base ordenada $(e_1, \ldots, e_n)$. Si definimos T mediante (2.8), conviene comprobar que T es lineal. Si $x = e_k$ para un cierto k, entonces todos los componentes de x son 0 excepto el k-ésimo, que es 1, con lo que (2.8) da $T(e_k) = u_k$, como queríamos.

Para demostrar que sólo existe una transformación lineal que satisface (2.7), sea T' otra y calculemos $T'(x)$. Encontramos que

$$T'(x) = T'\left(\sum_{k=1}^{n} x_k e_k\right) = \sum_{k=1}^{n} x_k T'(e_k) = \sum_{k=1}^{n} x_k u_k = T(x).$$

Puesto que $T'(x) = T(x)$ para todo x de V, tenemos $T' = T$, lo cual completa la demostración.

EJEMPLO. Determinar la transformación lineal $T: V_2 \to V_2$ que aplique los elementos base $i = (1, 0)$ y $j = (0, 1)$ del modo siguiente

$$T(i) = i + j, \qquad T(j) = 2i - j.$$

Solución. Si $x = x_1 i + x_2 j$ es un elemento arbitrario de V_2, entonces $T(x)$ viene dado por

$$T(x) = x_1 T(i) + x_2 T(j) = x_1(i + j) + x_2(2i - j) = (x_1 + 2x_2)i + (x_1 - x_2)j.$$

2.10 Representación matricial de las transformaciones lineales

El teorema 2.12 demuestra que una transformación lineal $T: V \to W$ de un espacio lineal de dimensión finita V está determinada por su acción sobre un conjunto dado de elementos base $e_1, \ldots, e_n$. Supongamos ahora que el espacio W también es de dimensión finita, por ejemplo $\dim W = m$, y sea $w_1, \ldots, w_n$ una base para W. (Las dimensiones n y m pueden ser o no iguales.) Puesto que T tiene los valores en W, cada elemento $T(e_k)$ puede expresarse, con unicidad, como una combinación lineal de los elementos de la base $w_1, \ldots, w_m$, por ejemplo

$$T(e_k) = \sum_{i=1}^{m} t_{ik} w_i ,$$

donde $t_{1k}, \ldots, t_{mk}$ son los componentes de $T(e_k)$ respecto a la base ordenada $(w_1, \ldots, w_m)$. Dispondremos verticalmente la m-pla $(t_{1k}, \ldots, t_{mk})$, como a continuación se indica:

(2.9)
$$\begin{bmatrix} t_{1k} \\ t_{2k} \\ \cdot \\ \cdot \\ \cdot \\ t_{mk} \end{bmatrix} .$$

Esto se llama *vector columna* o *matriz columna*. Tenemos una tal columna para cada uno de los n elementos $T(e_1), \ldots, T(e_n)$. Colocándolas una junto a otra y encerrándolas en un par de corchetes obtenemos la disposición rectangular siguiente:

$$\begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{bmatrix} .$$

Este cuadro se llama *matriz* y consta de m filas y n columnas. La llamamos matriz $m \times n$. La primera fila es la matriz $1 \times n$ $(t_{11}, t_{12}, \ldots, t_{1n})$. La matriz $m \times 1$ (2.9) es la k-ésima columna. Los escalares t_{ik} van afectados con dos índices, el

primero i indica la *fila* y el segundo k indica la *columna* en las cuales aparece t_{ik}. A t_{ik} le llamamos el *elemento* ik de la matriz. También se utiliza la notación abreviada

$$(t_{ik}), \qquad \text{o} \qquad (t_{ik})_{i,k=1}^{m,n},$$

para designar la matriz cuyo elemento ik es t_{ik}.

Así pues, toda transformación lineal T de un espacio n-dimensional V en un espacio m-dimensional W da origen a una matriz $m \times n$ (t_{ik}) cuyas columnas son los componentes de $T(e_1), \ldots, T(e_n)$ relativos a la base $(w_1, \ldots, w_m)$. La llamamos *representación matricial* de T relativa a unas bases ordenadas $(e_1, \ldots, e_n)$ de V y $(w_1, \ldots, w_m)$ para W. Una vez conocida la matriz (t_{ik}), los componentes de un elemento cualquiera $T(x)$ con relación a la base $(w_1, \ldots, w_m)$ pueden determinarse como se explica en el teorema que sigue.

TEOREMA 2.13. *Sea T una transformación lineal perteneciente a $\mathscr{L}(V, W)$, donde* $\dim V = n$ *y* $\dim W = m$. *Sean $(e_1, \ldots, e_n)$ y $(w_1, \ldots, w_m)$ bases ordenadas de V y W, respectivamente, y (t_{ik}) la matriz $m \times n$ cuyos elementos están determinados por las ecuaciones*

$$T(e_k) = \sum_{i=1}^{m} t_{ik} w_i, \quad \textit{para} \quad k = 1, 2, \ldots, n. \tag{2.11}$$

Entonces un elemento cualquiera

$$x = \sum_{k=1}^{n} x_k e_k \tag{2.10}$$

de V con componentes $(x_1, \ldots, x_n)$ relativo a $(e_1, \ldots, e_n)$ es aplicado por T en el elemento

$$T(x) = \sum_{i=1}^{m} y_i w_i \tag{2.12}$$

en W con componentes $(y_1, \ldots, y_m)$ relativos a $(w_1, \ldots, w_m)$. Los y_i están ligados a los componentes de x mediante las ecuaciones lineales

$$y_i = \sum_{k=1}^{n} t_{ik} x_k \quad \textit{para} \quad i = 1, 2, \ldots, m. \tag{2.13}$$

Demostración. Aplicando T a cada uno de los miembros de (2.11) y utilizando (2.10), obtenemos

$$T(x) = \sum_{k=1}^{n} x_k T(e_k) = \sum_{k=1}^{n} x_k \sum_{i=1}^{m} t_{ik} w_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} t_{ik} x_k \right) w_i = \sum_{i=1}^{m} y_i w_i ,$$

en donde cada y_i viene dada por (2.13). Esto completa la demostración.

Habiendo elegido un par de bases $(e_1, \ldots, e_n)$ y $(w_1, \ldots, w_m)$ para V y W, respectivamente, toda transformación lineal $T: V \to W$ tiene una representación matricial (t_{ik}). Recíprocamente, si disponemos de mn escalares colocados formando una matriz rectangular (t_{ik}) y elegimos un par de bases ordenadas para V y W, es fácil demostrar que existe exactamente una transformación lineal $T: V \to W$ que tiene esa representación matricial. Definimos T simplemente con los elementos base de V por medio de las ecuaciones (2.10). Entonces, según el teorema 2.12, existe una y sólo una transformación $T: V \to W$ con esos valores asignados. La imagen $T(x)$ de un punto x de V viene entonces dada por las ecuaciones (2.12) y (2.13).

EJEMPLO 1. *Construcción de una transformación lineal a partir de una matriz dada.* Supongamos que disponemos de la matriz 2×3.

$$\begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 4 \end{bmatrix}.$$

Elijamos las bases usuales de vectores coordenados unitarios para V_3 y V_2. Entonces la matriz dada representa una transformación lineal $T: V_3 \to V_2$ que aplica un vector cualquiera (x_1, x_2, x_3) de V_3 en el vector (y_1, y_2) de V_2 de acuerdo con las ecuaciones lineales

$$y_1 = 3x_1 + x_2 - 2x_3$$

$$y_2 = x_1 + 0x_2 + 4x_3 .$$

EJEMPLO 2. *Construcción de una representación matricial de una transformación lineal dada.* Sea V el espacio lineal de todos los polinomios reales $p(x)$ de grado ≤ 3. Este espacio tiene dimensión 4, y elegimos la base $(1, x, x^2, x^3)$. Sea D el operador derivación que aplica cada polinomio $p(x)$ de V en su derivada $p'(x)$. Podemos considerar D como una transformación lineal de V en W, donde W es el espacio tri dimensional de todos los polinomios reales de grado ≤ 2. En W elegimos la base $(1, x, x^2)$. Para encontrar la representación matricial de D relativa a esa elección de bases, transformamos (derivamos) cada elemento base

de V y lo expresamos como una combinación lineal de los elementos base de W. Así pues, encontramos que

$$D(1) = 0 = 0 + 0x + 0x^2, \qquad D(x) = 1 = 1 + 0x + 0x^2,$$

$$D(x^2) = 2x = 0 + 2x + 0x^2, \qquad D(x^3) = 3x^2 = 0 + 0x + 3x^2.$$

Los coeficientes de esos polinomios determinan las *columnas* de la representación matricial de D. Por consiguiente, la representación pedida viene dada por la siguiente matriz 3×4.

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Para hacer notar el hecho de que la representación matricial depende no solamente de los elementos base sino también de su orden, invirtamos el orden de los elementos base en W y utilicemos, en su lugar, la base ordenada $(x^2, x, 1)$. Entonces los elementos base de V se transforman en los mismos polinomios obtenidos antes, pero los componentes de éstos relativos a la nueva base $(x^2, x, 1)$ aparecen en orden inverso. Por consiguiente, la representación matricial de D ahora es

$$\begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Calculemos una tercera representación matricial de D, usando la base $(1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3)$ para V, y la base $(1, x, x^2)$ para W. Los elementos base de V se transforman así:

$$D(1) = 0, \qquad D(1 + x) = 1, \qquad D(1 + x + x^2) = 1 + 2x,$$

$$D(1 + x + x^2 + x^3) = 1 + 2x + 3x^2,$$

con lo que la representación matricial en este caso es

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

2.11 Construcción de una representación matricial en forma diagonal

Ya que es posible obtener distintas representaciones matriciales de una transformación lineal dada mediante la elección de bases distintas, parece natural intentar elegir bases de modo que la matriz resultante tenga una forma lo más sencilla posible. El teorema que sigue prueba que podemos hacer todos los elementos 0 excepto los de la diagonal que va desde el vértice superior izquierdo al inferior derecho. A lo largo de esa diagonal habrá una hilera de unos seguidos de ceros, siendo el número de unos igual al rango de la transformación. Una matriz (t_{ik}) con todos los elementos $t_{ik} = 0$ cuando $i \neq k$ se llama *matriz diagonal.*

TEOREMA 2.14. *Sean* V *y* W *espacios lineales de dimensión finita, con* $\dim V = n$ *y* $\dim W = m$. *Supongamos que* $T \in \mathscr{L}(V, W)$ *y que* $r = \dim T(V)$ *represente el rango de* T. *Existen entonces una base* $(e_1, \ldots, e_n)$ *para* V *y otra* $(w_1, \ldots, w_n)$ *para* W *tales que*

$$(2.14) \qquad T(e_i) = w_i \quad \textit{para} \quad i = 1, 2, \ldots, r,$$

y

$$(2.15) \qquad T(e_i) = O \quad \textit{para} \quad i = r + 1, \ldots, n.$$

Por consiguiente, la matriz (t_{ik}) *de* T *relativa a esas bases tiene todos los elementos cero excepto los* r *elementos de la diagonal que valen*

$$t_{11} = t_{22} = \cdots = t_{rr} = 1.$$

Demostración. Construimos primero una base para W. Puesto que $T(V)$ es un subespacio de W con $\dim T(V) = r$, el espacio $T(V)$ tiene una base de r elementos en W, sean éstos $w_1, \ldots, w_r$. Según el teorema 1.7, esos elementos forman un subconjunto de una cierta base para W. Por consiguiente podemos adjuntar unos elementos $w_{r+1}, \ldots, w_m$ de modo que

$$(2.16) \qquad (w_1, \ldots, w_r, w_{r+1}, \ldots, w_m)$$

sea una base para W.

Seguidamente construimos una base para V. Cada uno de los r primeros elementos w_i de (2.16) es la imagen por lo menos de un elemento de V. Elijamos uno de tales elementos de V y llamémosle e_i. Entonces $T(e_i) = w_i$ para $i = 1, 2, \ldots, r$ así que (2.14) se satisface. Sea ahora k la dimensión del núcleo $N(T)$. Según el teorema 2.3 tenemos $n = k + r$. Puesto que $N(T) = k$, el espacio $N(T)$

tiene una base que consta de k elementos de V que designamos por $e_{r+1}, \ldots, e_{r+k}$. Para cada uno de esos elementos, la ecuación (2.15) se satisface. Por lo tanto, para completar la demostración, tenemos que demostrar que el conjunto ordenado

$$(e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+k}) \tag{2.17}$$

es una base para V. Ya que $\dim V = n = r + k$, sólo tenemos que demostrar que esos elementos son independientes. Supongamos que una cierta combinación lineal de ellos sea cero, por ejemplo

$$\sum_{i=1}^{r+k} c_i e_i = O. \tag{2.18}$$

Aplicando T y haciendo uso de las ecuaciones (2.14) y (2.15), encontramos que

$$\sum_{i=1}^{r+k} c_i T(e_i) = \sum_{i=1}^{r} c_i w_i = O.$$

Pero $w_1, \ldots, w_r$ son independientes, y por tanto $c_1 = \ldots = c_r = 0$. Por consiguiente, los r primeros términos de (2.18) son cero, por lo cual (2.18) se reduce a

$$\sum_{i=r+1}^{r+k} c_i e_i = O.$$

Pero $e_{r+1}, \ldots, e_{r+k}$ son independientes puesto que forman una base para $N(T)$, y por tanto $c_{r+1} = \ldots = c_{r+k} = 0$. Por consiguiente, todos los c_i de (2.18) son cero, luego los elementos de (2.17) forman una base para V. Esto completa la demostración.

EJEMPLO. Nos referimos al ejemplo 2 de la sección 2.10, donde D es el operador derivación que aplica el espacio V de los polinomios de grado ≤ 3 en el espacio W de los polinomios de grado ≤ 2. En este ejemplo, el recorrido $T(V) = W$, así que T tiene rango 3. Aplicando el método seguido en el teorema 2.14, elegimos cualquier base para W, por ejemplo la base $(1, x, x^2)$. Un conjunto de polinomios de V que se aplica sobre esos elementos es $(x, \frac{1}{2}x^2, \frac{1}{3}x^3)$. Ampliamos este conjunto para lograr una base para V adjuntando el polinomio constante 1, que es una base para el núcleo de D. Por consiguiente, si empleamos la base $(x, \frac{1}{2}x^2, \frac{1}{3}x^3, 1)$ para V y la base $(1, x, x^2)$ para W, la correspondiente representación matricial para D tiene la forma diagonal

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

2.12 Ejercicios

En todos los ejercicios en los que se considere el espacio vectorial V_n, la base que se utilizará será la de los vectores coordenados unitarios si no se dice lo contrario. En los ejercicios relativos a la matriz de una transformación lineal $T: V \to W$ siendo $V = W$, si no se indica lo contrario tomaremos la misma base en V y en W.

1. Determinar la matriz de cada una de las siguientes transformaciones lineales de V_n en V_n:
   a) la transformación idéntica,
   b) la transformación cero,
   c) multiplicación por un escalar fijo c.
2. Determinar la matriz para cada una de las siguientes proyecciones.

   a) $T: V_3 \to V_2$, donde $T(x_1, x_2, x_3) = (x_1, x_2)$.
   b) $T: V_3 \to V_2$, donde $T(x_1, x_2, x_3) = (x_2, x_3)$.
   c) $T: V_5 \to V_3$, donde $T(x_1, x_2, x_3, x_4, x_5) = (x_2, x_3, x_4)$.

3. Una transformación lineal $T: V_2 \to V_2$ aplica los vectores base $\boldsymbol{i}$ y $\boldsymbol{j}$ como sigue:

$$T(\boldsymbol{i}) = \boldsymbol{i} + \boldsymbol{j}, \qquad T(\boldsymbol{j}) = 2\boldsymbol{i} - \boldsymbol{j}.$$

   a) Calcular $T(3\boldsymbol{i} - 4\boldsymbol{j})$ y $T^2(3\boldsymbol{i} - 4\boldsymbol{j})$ en función de $\boldsymbol{i}$ y $\boldsymbol{j}$.
   b) Determinar la matriz de T y de T^2.
   c) Resolver la parte b) si la base $(\boldsymbol{i}, \boldsymbol{j})$ se reemplaza por (e_1, e_2), siendo $e_1 = \boldsymbol{i} - \boldsymbol{j}$, $e_2 = 3\boldsymbol{i} + \boldsymbol{j}$.
4. Una transformación lineal $T: V_2 \to V_2$ se define así: Cada vector (x, y) se transforma en su simétrico respecto al eje y y luego se duplica su longitud para obtener $T(x, y)$. Determinar la matriz de T y la de T^2.
5. Sea $T: V_3 \to V_3$ una transformación lineal tal que

$$T(\boldsymbol{k}) = 2\boldsymbol{i} + 3\boldsymbol{j} + 5\boldsymbol{k}, \qquad T(\boldsymbol{j} + \boldsymbol{k}) = \boldsymbol{i}, \qquad T(\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}) = \boldsymbol{j} - \boldsymbol{k}.$$

   a) Calcular $T(\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k})$ y determinar la dimensión del núcleo y el rango de T.
   b) Determinar la matriz de T.
6. Para la transformación lineal del ejercicio 5, se consideran las dos bases coincidentes con (e_1, e_2, e_3), siendo $e_1 = (2, 3, 5)$, $e_2 = (1, 0, 0)$, $e_3 = (0, 1, -1)$ y determinar la matriz T relativa a las nuevas bases.
7. Una transformación lineal $T: V_3 \to V_2$ aplica los vectores base como sigue: $T(\boldsymbol{i}) = (0, 0)$, $T(\boldsymbol{j}) = (1, 1)$, $T(\boldsymbol{k}) = (1, -1)$.
   a) Calcular $T(4\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k})$ y determinar la dimensión del núcleo y el rango de T.
   b) Determinar la matriz de T.
   c) Utilizando la base $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ en V_3 y la (w_1, w_2) en V_2, siendo $w_1 = (1, 1)$, $w_2 = (1, 2)$, determinar la matriz de T relativa a esas bases.
   d) Hallar las bases (e_1, e_2, e_3) para V_3 y (w_1, w_2) para V_2 para las cuales la matriz de T tenga la forma diagonal.
8. Una transformación lineal $T: V_2 \to V_3$ aplica los vectores base como sigue: $T(\boldsymbol{i}) = (1, 0, 1)$, $T(\boldsymbol{j}) = (-1, 0, 1)$.
   a) Calcular $T(2\boldsymbol{i} - 3\boldsymbol{j})$ y determinar la dimensión del núcleo y el rango de T.
   b) Determinar la matriz de T.
   c) Hallar bases (e_1, e_2) para V_2 y $(w_1 . w_2, w_3)$ para V_3 para las cuales la matriz de T tiene forma diagonal.
9. Resolver el ejercicio 8 si $T(\boldsymbol{i}) = (1, 0, 1)$ y $T(\boldsymbol{j}) = (1, 1, 1)$.

10. Sean V y W dos espacios lineales, ambos de dimensión 2 y con la misma base (e_1, e_2). Sea $T: V \to W$ una transformación lineal tal que

$$T(e_1 + e_2) = 3e_1 + 9e_2, \qquad T(3e_1 + 2e_2) = 7e_1 + 23e_2.$$

a) Calcular $T(e_2 - e_1)$ y determinar la dimensión del núcleo y el rango de T.
b) Determinar la matriz de T relativa a la base dada.
c) Utilizar para V la base (e_1, e_2) y hallar una nueva base de la forma $(e_1 + ae_2, 2e_1 + be_2)$ para W, para la que la matriz de T tenga la forma diagonal.

En el espacio lineal de todas las funciones reales, cada uno de los siguientes conjuntos es independiente y genera un subespacio V de dimensión finita. Utilizar el conjunto dado como base para V y sea $D: V \to V$ el operador derivación. En cada caso, hallar la matriz de D y la de D^2 relativa a la base que se elige.

11. $(\operatorname{sen} x, \cos x)$.
12. $(1, x, e^x)$.
13. $(1, 1 + x, 1 + x + e^x)$.
14. (e^x, xe^x).
15. $(-\cos x, \operatorname{sen} x)$.
16. $(\operatorname{sen} x, \cos x, x \operatorname{sen} x, x \cos x)$.
17. $(e^x \operatorname{sen} x, e^x \cos x)$.
18. $(e^{2x} \operatorname{sen} 3x, e^{2x} \cos 3x)$.
19. Elegir la base $(1, x, x^2, x^3)$ en el espacio lineal V de todos los polinomios reales de grado ≤ 3. Sean D el operador derivación y $T: V \to V$ la transformación lineal que aplica $p(x)$ en $xp'(x)$. Con relación a la base dada, determinar la matriz de cada una de las transformaciones siguientes: a) T; b) DT; c) TD; d) $TD - DT$; e) T^2; f) $T^2D^2 - D^2T^2$.
20. Con respecto al ejercicio 19. Sea W la imagen de V a través de TD. Hallar bases para V y W para las que la matriz TD tenga forma diagonal.

2.13 Espacios lineales de matrices

Hemos visto cómo las matrices se presentan espontáneamente como representaciones de las transformaciones lineales. También se pueden considerar las matrices como elementos existentes con independencia de las transformaciones lineales. Como tales elementos, forman otra clase de objetos matemáticos que pueden definirse por medio de las operaciones algebraicas que pueden realizarse con ellos. La relación con las transformaciones lineales da origen a esas definiciones, pero tal relación será por el momento ignorada.

Sean m y n dos enteros positivos y sea $I_{m,n}$ el conjunto de todos los pares de enteros (i, j) tales que $1 \leq i \leq m$, $1 \leq j \leq n$. Cualquier función A cuyo dominio sea $I_{m,n}$ se denomina *matriz* $m \times n$. El valor de la función $A(i, j)$ se llama *elemento ij* de la matriz y se designará también por a_{ij}. Ordinariamente se disponen todos los valores de la función en un rectángulo que consta de m filas y n columnas, del modo siguiente

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Los elementos a_{ij} pueden ser objetos arbitrarios de naturaleza cualquiera. Normalmente serán números reales o complejos, pero a veces conviene considerar matrices cuyos elementos son otros objetos, por ejemplo, funciones. También designaremos las matrices mediante la notación abreviada

$$A = (a_{ij})_{i,j=1}^{m,n} \qquad \text{o} \qquad A = (a_{ij})\,.$$

Si $m = n$, la matriz se llama *cuadrada*. Una matriz $1 \times n$ se llama *matriz fila*; una matriz $m \times 1$ es una *matriz columna*.

Dos funciones son iguales si y sólo si tienen el mismo dominio y toman los mismos valores en cada elemento del dominio. Puesto que las matrices son funciones, dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ son iguales si y sólo si tienen el mismo número de filas, el mismo número de columnas, e iguales elementos $a_{ij} = b_{ij}$ para cada par (i, j).

Supongamos ahora que los elementos son números (reales o complejos) y definamos la adición de matrices y la multiplicación por escalares siguiendo el mismo método que para funciones reales o complejas cualesquiera.

DEFINICIÓN. *Si $A = (a_{ij})$ y $B = (b_{ij})$ son dos matrices $m \times n$ y si c es un escalar cualquiera, definimos las matrices $A + B$ y cA del modo siguiente*

$$A + B = (a_{ij} + b_{ij})\,, \qquad cA = (ca_{ij})\,.$$

La suma sólo se define cuando A y B tienen el mismo tamaño $m \times n$.

EJEMPLO. Si

$$A = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 0 & 4 \end{bmatrix} \qquad \text{y} \qquad B = \begin{bmatrix} 5 & 0 & 1 \\ 1 & -2 & 3 \end{bmatrix},$$

tenemos entonces

$$A + B = \begin{bmatrix} 6 & 2 & -2 \\ 0 & -2 & 7 \end{bmatrix}, \qquad 2A = \begin{bmatrix} 2 & 4 & -6 \\ -2 & 0 & 8 \end{bmatrix}, \qquad (-1)B = \begin{bmatrix} -5 & 0 & -1 \\ -1 & 2 & -3 \end{bmatrix}.$$

Definimos la matriz O como la matriz $m \times n$ cuyos elementos son todos 0. Con esas definiciones, es inmediato el ejercicio de comprobar que el conjunto de todas las matrices $m \times n$ es un espacio lineal. Lo designamos con $M_{m,n}$. Si los elementos son números reales, el espacio $M_{m,n}$ es un espacio lineal real. Si son números complejos, $M_{m,n}$ es un espacio lineal complejo. Es también fácil demostrar

que este espacio es de dimensión $m \times n$. En efecto, una base para $M_{m,n}$ consta de mn matrices que tienen un elemento igual a 1 y todos los demás iguales a 0. Por ejemplo, las seis matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

forman una base para el conjunto de todas las matrices 2×3.

2.14 Isomorfismo entre transformaciones lineales y matrices

Volvamos ahora a la relación entre matrices y transformaciones lineales. Sean V y W dos espacios lineales de dimensión finita con $\dim V = n$ y $\dim W = m$. Elijamos una base $(e_1, \ldots, e_n)$ para V y otra $(w_1, \ldots, w_m)$ para W. En esta discusión esas bases se mantienen fijas. Designemos con $\mathscr{L}(V, W)$ el espacio lineal de todas las transformaciones lineales de V en W. Si $T \in \mathscr{L}(V, W)$, sea $m(T)$ la matriz de T relativa a las bases dadas. Recordemos que $m(T)$ se define como sigue.

La imagen de cada elemento base e_k se expresa como una combinación lineal de los elementos base de W:

$$T(e_k) = \sum_{i=1}^{m} t_{ik} w_i \quad \text{para} \quad k = 1, 2, \ldots, n. \tag{2.19}$$

Los multiplicadores escalares t_{ik} son los elementos ik de $m(T)$. Así pues, tenemos

$$m(T) = (t_{ik})_{i,k=1}^{m,n}. \tag{2.20}$$

La ecuación (2.20) define una nueva función m cuyo dominio es $\mathscr{L}(V, W)$ y cuyos valores son matrices de $M_{m,n}$. Puesto que toda matriz $m \times n$ es la matriz $m(T)$ para una cierta T de $\mathscr{L}(V, W)$, el recorrido de m es $M_{m,n}$. El teorema siguiente prueba que la transformación $m: \mathscr{L}(V, W) \to M_{m,n}$ es lineal y uno a uno en $\mathscr{L}(V, W)$.

TEOREMA 2.15. TEOREMA DE ISOMORFISMO. *Para cualesquiera S y T de $\mathscr{L}(V, W)$ y todos los escalares c, tenemos*

$$m(S + T) = m(S) + m(T) \qquad y \qquad m(cT) = cm(T).$$

Además,

$$m(S) = m(T) \quad \textit{implica} \quad S = T,$$

así que m es uno a uno en $\mathscr{L}(V, W)$.

Demostración. La matriz $m(T)$ está formada con los factores t_{ik} de (2.19). Del mismo modo, la matriz $m(S)$ está constituida con los factores s_{ik} de las ecuaciones

(2.21) $$S(e_k) = \sum_{i=1}^{m} s_{ik}w_i \quad \text{para} \quad k = 1, 2, \ldots, n .$$

Puesto que tenemos

$$(S + T)(e_k) = \sum_{i=1}^{m} (s_{ik} + t_{ik})w_i \qquad \text{y} \qquad (cT)(e_k) = \sum_{i=1}^{m} (ct_{ik})w_i ,$$

obtenemos $m(S + T) = (s_{ik} + t_{ik}) = m(S) + m(T)$ y $m(cT) = (ct_{ik}) = cm(T)$. Esto demuestra que m es lineal.

Para demostrar que m es uno a uno, supongamos que $m(S) = m(T)$, siendo $S = (s_{ik})$ y $T = (t_{ik})$. Las ecuaciones (2.19) y (2.21) demuestran que $S(e_k) = T(e_k)$ para cada elemento base e_k, así que $S(x) = T(x)$ para todo x de V, y por tanto $S = T$.

Observación: La función m es un *isomorfismo.* Elegidas unas bases, m establece una correspondencia uno a uno entre el conjunto de las transformaciones $\mathscr{L}(V, W)$ y el conjunto $M_{m,n}$ de las matrices $m \times n$. Las operaciones de adición y multiplicación por escalares se conservan a través de esa correspondencia. Los espacios lineales $\mathscr{L}(V, W)$ y $M_{m,n}$ se dice que son *isomorfos* Incidentalmente, el teorema 2.11 demuestra que el dominio de una transformación lineal uno a uno tiene la dimensión igual a su recorrido. Por consiguiente, $\dim \mathscr{L}(V, W) = \dim M_{m,n} = mn$.

Si $V = W$ y elegimos la misma base para ambos, la matriz $m(I)$ correspondiente a la transformación idéntica $I: V \to V$ es una matriz diagonal con los elementos de la diagonal iguales a 1 y todos los demás iguales a 0. Esta se llama *identidad* o *matriz unidad* y se designa con I o con I_n.

2.15 Multiplicación de matrices

Algunas transformaciones lineales pueden multiplicarse por medio de la composición. Definiremos ahora la multiplicación de matrices de manera que el producto de dos matrices corresponda a la composición de las transformaciones lineales que ellos representan.

Recordemos que si $T: U \to V$ y $S: V \to W$ son transformaciones lineales, su composición $ST: U \to W$ es una transformación lineal dada por

$$ST(x) = S[T(x)] \quad \text{para todo } x \text{ de } U .$$

Supongamos que U, V, y W son de dimensión finita, por ejemplo

$$\dim U = n , \qquad \dim V = p , \qquad \dim W = m .$$

Elijamos bases para U, V, y W. Con relación a esas bases $m(S)$ es una matriz $m \times p$, T es una matriz $p \times n$, y ST es una matriz $m \times n$. La siguiente definición de multiplicación de matrices nos permite deducir la relación $m(ST) = m(S)m(T)$. Esto extiende a los productos la propiedad de isomorfismo.

DEFINICIÓN. *Sean A una matriz $m \times p$ cualquiera, y B una matriz $p \times n$ cualquiera, tales como*

$$A = (a_{ij})_{i,j=1}^{m,p} \qquad \text{y} \qquad B = (b_{ij})_{i,j=1}^{p,n}.$$

El producto AB se define como la matriz $m \times n$, $C=(c_{ij})$, cuyo elemento ij viene dado por

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}. \tag{2.22}$$

Observación: El producto AB sólo está definido si el número de columnas de A es igual al de filas de B.

Si escribimos A_i para expresar la fila i de A y B^j para la columna j de B, y las imaginamos como vectores de dimensión p, la suma (2.22) es simplemente el producto escalar $A_i \cdot B^j$. Es decir, el elemento ij de AB es el producto escalar de la fila i de A por la columna j de B:

$$AB = (A_i \cdot B^j)_{i,j=1}^{m,n}.$$

Así pues, la multiplicación de matrices puede considerarse como una generalización del producto escalar.

EJEMPLO 1. Sean $A = \begin{bmatrix} 3 & 1 & 2 \\ -1 & 1 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 6 \\ 5 & -1 \\ 0 & 2 \end{bmatrix}$ Puesto que A es 2×3 y B es 3×2, el producto AB es la matriz 2×2

$$AB = \begin{bmatrix} A_1 \cdot B^1 & A_1 \cdot B^2 \\ A_2 \cdot B^1 & A_2 \cdot B^2 \end{bmatrix} = \begin{bmatrix} 17 & 21 \\ 1 & -7 \end{bmatrix}.$$

Los elementos de AB se calculan así

$$A_1 \cdot B^1 = 3 \cdot 4 + 1 \cdot 5 + 2 \cdot 0 = 17, \qquad A_1 \cdot B^2 = 3 \cdot 6 + 1 \cdot (-1) + 2 \cdot 2 = 21,$$

$$A_2 \cdot B^1 = (-1) \cdot 4 + 1 \cdot 5 + 0 \cdot 0 = 1, \quad A_2 \cdot B^2 = (-1) \cdot 6 + 1 \cdot (-1) + 0 \cdot 2 = -7.$$

EJEMPLO 2. Sea

$$A = \begin{bmatrix} 2 & 1 & -3 \\ 1 & 2 & 4 \end{bmatrix} \qquad \text{y} \qquad B = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}.$$

Aquí A es 2×3 y B es 3×1, con lo que AB es la matriz 2×1 dada por

$$AB = \begin{bmatrix} A_1 \cdot B^1 \\ A_2 \cdot B^1 \end{bmatrix} = \begin{bmatrix} -9 \\ 8 \end{bmatrix},$$

Puesto que $A_1 \cdot B^1 = 2 \cdot (-2) + 1 \cdot 1 + (-3) \cdot 2 = -9$ y $A_2 \cdot B^1 = 1 \cdot (-2) + 2 \cdot 1 + 4 \cdot 2 = 8$.

EJEMPLO 3. Si A y B son dos matrices cuadradas del mismo tamaño, entonces AB y BA están definidas. Por ejemplo, si

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \qquad \text{y} \qquad B = \begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix},$$

encontramos que

$$AB = \begin{bmatrix} 13 & 8 \\ 2 & -2 \end{bmatrix}, \qquad BA = \begin{bmatrix} -1 & 10 \\ 3 & 12 \end{bmatrix}.$$

Este ejemplo prueba que en general $AB \neq BA$. Si $AB = BA$, decimos que A y B son *permutables* o que *conmutan.*

EJEMPLO 4. Si I_p es la matriz identidad $p \times p$, entonces $I_pA = A$ para toda matriz A, $p \times n$, y $BI_p = B$ para toda matriz B, $m \times p$. Por ejemplo:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

Demostramos seguidamente que la matriz de una composición ST es el producto de las matrices $m(S)$ y $m(T)$.

TEOREMA 2.16. *Sean $T:U \to V$ y $S:V \to W$ dos transformaciones lineales, donde U, V, W son espacios lineales de dimensión finita. Entonces, elegidas unas bases fijas, las matrices de S, T y ST están relacionadas por la ecuación*

$$m(ST) = m(S)m(T).$$

Demostración. Supongamos que $\dim U = n$, $\dim V = p$, $\dim W = m$. Sean $(u_1, \ldots, u_n)$ una base para U, $(v_1, \ldots, v_p)$ una base para V, y $(w_1, \ldots, w_m)$ una base para W. Con relación a esas bases tenemos

$$m(S) = (s_{ij})_{i,j=1}^{m,p}, \qquad \text{donde} \quad S(v_k) = \sum_{i=1}^{m} s_{ik}w_i \qquad \text{para} \quad k = 1, 2, \ldots, p,$$

y

$$m(T) = (t_{ij})_{i,j=1}^{p,n}, \qquad \text{donde} \quad T(u_j) = \sum_{k=1}^{p} t_{kj}v_k \qquad \text{para} \quad j = 1, 2, \ldots, n.$$

Por consiguiente, tenemos

$$ST(u_j) = S[T(u_j)] = \sum_{k=1}^{p} t_{kj}S(v_k) = \sum_{k=1}^{p} t_{kj} \sum_{i=1}^{m} s_{ik}w_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{p} s_{ik}t_{kj}\right) w_i,$$

con ello encontramos que

$$m(ST) = \left(\sum_{k=1}^{p} s_{ik}t_{kj}\right)_{i,j=1}^{m,n} = m(S)m(T).$$

Ya hemos observado que la multiplicación de matrices no siempre satisface la ley conmutativa. El teorema siguiente prueba que satisface las leyes asociativa y distributiva.

TEOREMA 2.17. LEYES ASOCIATIVA Y DISTRIBUTIVA PARA LA MULTIPLICACIÓN DE MATRICES. *Dadas las matrices A, B, C.*

a) *Si los productos $A(BC)$ y $(AB)C$ tienen sentido, tenemos*

$$A(BC) = (AB)C \qquad (\textit{ley asociativa}).$$

b) *Supongamos que A y B sean del mismo tamaño. Si AC y BC tienen sentido, tenemos*

$$(A + B)C = AC + BC \qquad (\textit{ley distributiva por la derecha}),$$

en tanto que si CA y CB tienen sentido, tenemos

$$C(A + B) = CA + CB \qquad (\textit{ley distributiva por la izquierda}).$$

Demostración. Esas propiedades pueden deducirse directamente a partir de la definición de multiplicación de matrices, pero preferimos razonar del siguiente modo. Introduzcamos los espacios lineales de dimensión finita U, V, W, X y las transformaciones lineales $T: U \to V$, $S: V \to W$, $R: W \to X$ tales que fijadas unas bases, tenemos

$$A = m(R), \qquad B = m(S), \qquad C = m(T).$$

Según el teorema 2,16, es $m(RS) = AB$ y $m(ST) = BC$. De la ley asociativa para la composición, encontramos que $R(ST) = (RS)T$. Aplicando el teorema 2.16 una vez más a esa ecuación, obtenemos $m(R)m(ST) = m(RS)m(T)$ o $A(BC) = (AB)C$, que demuestra a). La demostración de b) puede hacerse con un razonamiento parecido.

DEFINICIÓN. *Si A es una matriz cuadrada, definimos la potencia entera de A por inducción como sigue*:

$$A^0 = I, \qquad A^n = AA^{n-1} \quad \text{para } n \geq 1.$$

2.16 Ejercicios

1. Si $A = \begin{bmatrix} 1 & -4 & 2 \\ -1 & 4 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 5 & -2 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 2 \\ 1 & -1 \\ 1 & -3 \end{bmatrix}$, calcular $B + C$, AB, BA, AC, CA, $A(2B - 3C)$.

2. Sea $A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$. Hallar todas las matrices B, 2×2, tales que a) $AB = O$; b) $BA = O$.

3. Hallar en cada caso a, b, c, d para que se satisfaga la ecuación dada.

$$\text{a)} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 6 \\ 5 \end{bmatrix}; \qquad \text{b)} \begin{bmatrix} a & b & c & d \\ 1 & 4 & 9 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 6 & 6 \\ 1 & 9 & 8 & 4 \end{bmatrix}.$$

4. Calcular en cada caso $AB - BA$.

a) $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 1 & 1 \\ -4 & 2 & 0 \\ 1 & 2 & 1 \end{bmatrix};$

b) $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 & -2 \\ 3 & -2 & 4 \\ -3 & 5 & 11 \end{bmatrix}.$

5. Si A es una matriz cuadrada, demostrar que $A^nA^m = A^{n+m}$ para todos los enteros $m \geq 0$, $n \geq 0$.

6. Sea $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Comprobar que $A^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ y calcular A^n.

7. Sea $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$. Comprobar que $A^2 = \begin{bmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{bmatrix}$ y calcular A^n.

8. Sea $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Comprobar que $A^2 = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Calcular A^3 y A^4. Suponer una fórmula general para A^n y demostrarla por inducción.

9. Sea $A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$. Demostrar que $A^2 = 2A - I$ y calcular A^{100}.

10. Hallar todas las matrices A, 2×2, tales que $A^2 = O$.
11. a) Probar que una matriz A, 2×2, conmuta con cualquier matriz 2×2 si y sólo si A conmuta con cada una de las cuatro matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

b) Hallar todas esas matrices A.
12. La ecuación $A^2 = I$ se satisface para cada una de las matrices 2×2

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ c & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & b \\ 0 & -1 \end{bmatrix},$$

donde b y c son números reales arbitrarios. Hallar todas las matrices A, 2×2, tales que $A^2 = I$.

13. Si $A = \begin{bmatrix} 2 & -1 \\ -2 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix}$, hallar matrices C y D, 2×2, tales que $AC = B$ y $DA = B$.

14. a) Comprobar que las identidades algebraicas

$$(A + B)^2 = A^2 + 2AB + B^2 \qquad \text{y} \qquad (A + B)(A - B) = A^2 - B^2$$

no son ciertas para las matrices 2×2, $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$.

b) Modificar el segundo miembro de esas identidades para obtener fórmulas válidas para todas las matrices cuadradas A y B.
c) ¿Para qué matrices A y B son válidas las identidades establecidas en a)?

2.17 Sistemas de ecuaciones lineales

Sea $A = (a_{ij})$ una matriz $m \times n$ de números dada, y sean $c_1, \ldots, c_m$ otros m números. Un conjunto de m ecuaciones de la forma

$$\sum_{k=1}^{n} a_{ik} x_k = c_i \quad \text{para} \quad i = 1, 2, \ldots, m, \tag{2.23}$$

se llama sistema de m ecuaciones lineales con n incógnitas. Consideramos $x_1, \ldots, x_n$ como incógnitas. Una *solución* del sistema es una n-pla cualquiera de números $(x_1, \ldots, x_n)$ para los que se satisfacen todas las ecuaciones. La matriz A se llama *matriz de los coeficientes* del sistema.

Los sistemas lineales pueden estudiarse por medio de las transformaciones lineales. Elegimos las bases usuales de vectores coordenados unitarios en V_n y V_m. La matriz de los coeficientes A determina una transformación lineal, $T: V_n \to V_m$, que aplica un vector arbitrario $x = (x_1, \ldots, x_n)$ de V_n en el vector $y = (y_1, \ldots, y_m)$ de V_m dado por las ecuaciones lineales

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad \text{para} \quad i = 1, 2, \ldots, m.$$

Sea $c = (c_1, \ldots, c_m)$ el vector de V_m cuyos componentes son los números que aparecen en el sistema (2.23). Este sistema puede escribirse más sencillamente poniendo

$$T(x) = c\,.$$

El sistema tiene una solución si y sólo si c está en el recorrido de T. Si un solo x de V_n se aplica en c, el sistema tiene una sola solución. Si más de un x se aplica en c, el sistema admite más de una solución.

EJEMPLO 1. *Un sistema sin solución.* El sistema $x + y = 1$, $x + y = 2$ no tiene solución. La suma de dos números no puede ser a la vez 1 y 2.

EJEMPLO 2. *Un sistema con solución única.* El sistema $x + y = 1$, $x - y = 0$ tiene exactamente una solución: $(x, y) = (\frac{1}{2}, \frac{1}{2})$.

EJEMPLO 3. *Un sistema con más de una solución.* El sistema $x + y = 1$, que consta de una ecuación con dos incógnitas, tiene más de una solución. Dos números cualesquiera cuya suma sea 1 dan una solución.

A cada sistema lineal (2.23), podemos asociar otro sistema

$$\sum_{k=1}^{n} a_{ik} x_k = 0 \qquad \text{para} \quad i = 1, 2, \ldots, m\,,$$

obtenido reemplazando cada c_i en (2.23) por 0. Éste se llama el *sistema homogéneo* correspondiente al (2.23). Si $c \neq O$, el sistema (2.23) se llama *no homogéneo.* Un vector x de V_n satisfará el sistema homogéneo si y sólo si

$$T(x) = O\,,$$

donde T es la transformación lineal determinada por la matriz de los coeficientes. El sistema homogéneo tiene siempre la solución $x = O$, pero puede tener otras. El conjunto de soluciones del sistema homogéneo es el núcleo de T. El teorema siguiente expone la relación entre las soluciones del sistema homogéneo y las del sistema no homogéneo.

TEOREMA 2.18. *Supongamos que el sistema no homogéneo* (2.23) *tenga una solución, por ejemplo b.*

a) *Si un vector x es una solución del sistema no homogéneo, entonces el vector $v = x - b$ es una solución del correspondiente sistema homogéneo.*

b) *Si un vector v es una solución del sistema homogéneo, el vector $x = v + b$ es una solución del sistema no homogéneo.*

Demostración. Sea $T: V_n \to V_m$ la transformación lineal determinada por la matriz de los coeficientes, como antes se ha dicho. Puesto que b es una solución del sistema no homogéneo tenemos $T(b) = c$. Sean x y v dos vectores de V_n tales que $v = x - b$. Entonces tenemos

$$T(v) = T(x - b) = T(x) - T(b) = T(x) - c$$

Por consiguiente $T(x) = c$ si y sólo si $T(v) = O$. Esto demuestra a la vez a) y b).

Este teorema prueba que el problema de hallar todas las soluciones de un sistema no homogéneo se escinde en dos partes: 1) Hallar todas las soluciones v del sistema homogéneo, esto es, determinando el núcleo de T; y 2) hallar una solución particular b del sistema no homogéneo. Sumando b a cada uno de los vectores v del núcleo T, se obtienen todas las soluciones $x = v + b$ del sistema no homogéneo.

Sea k la dimensión de $N(T)$. Si podemos encontrar k soluciones *independientes* $v_1, \ldots, v_k$ del sistema homogéneo, ellas formarán una base para $N(T)$, y podemos obtener cualquier v de $N(T)$ formando todas las combinaciones lineales

$$v = t_1 v_1 + \cdots + t_k v_k ,$$

donde $t_1, \ldots, t_k$ son escalares arbitrarios. Esta combinación lineal se llama *solución general del sistema homogéneo.* Si b es una solución particular del sistema no homogéneo, entonces todas las soluciones x vienen dadas por

$$x = b + t_1 v_1 + \cdots + t_k v_k .$$

Esta combinación lineal se llama *solución general del sistema no homogéneo.* El teorema 2.18 puede ponerse en esta otra forma:

TEOREMA 2.19. *Sea $T: V_n \to V_m$ la transformación lineal tal que $T(x) = y$, donde $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_m)$, e*

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad para \quad i = 1, 2, \ldots, m .$$

Sea k la dimensión del núcleo de T. Si $v_1, \ldots, v_k$ *son k soluciones independientes del sistema homogéneo* $T(x) = O$, *y si b es una solución particular del sistema no homogéneo* $T(x) = c$, *entonces la solución general del sistema no homogéneo es*

$$x = b + t_1 v_1 + \cdots + t_k v_k ,$$

siendo $t_1, \ldots, t_k$ *escalares cualesquiera.*

Este teorema no nos dice cómo encontrar una solución particular b del sistema no homogéneo, ni las soluciones $v_1, \ldots, v_k$ del sistema homogéneo. Nos dice tan sólo lo que puede obtenerse cuando el sistema no homogéneo tenga una solución. El siguiente ejemplo, aunque muy sencillo, ilustra el teorema.

EJEMPLO. El sistema $x + y = 2$ tiene como sistema homogéneo asociado la ecuación $x + y = 0$. Por consiguiente, el núcleo consta de todos los vectores de V_2 de la forma $(t, -t)$, siendo t arbitrario. Puesto que $(t, -t) = t(1, -1)$, éste es un subespacio uni dimensional de V_2 con base $(1, -1)$. Una solución particular del sistema no homogéneo es $(0, 2)$. Por tanto, la solución general del sistema no homogéneo viene dada por

$$(x, y) = (0, 2) + t(1, -1) \qquad \text{o} \qquad x = t, \qquad y = 2 - t ,$$

siendo t arbitrario.

2.18 Técnicas de cálculo

Volvamos al problema del cálculo efectivo de las soluciones de un sistema lineal no homogéneo. Aunque se han desarrollado muchos métodos para atacar este problema, todos exigen cálculos considerables si el sistema es de gran tamaño. Por ejemplo, para resolver un sistema de diez ecuaciones con el mismo número de incógnitas pueden ser necesarias varias horas de cálculos, incluso con la ayuda de un calculador manual.

Vamos a comentar un método muy utilizado, que se llama *método de eliminación de Gauss-Jordan,* que es relativamente sencillo y puede programarse fácilmente para calculadores electrónicos de alta velocidad. El método consiste en la aplicación de tres operaciones fundamentales a las ecuaciones lineales del sistema:

1) *Intercambio de dos ecuaciones.*
2) *Multiplicación de todos los términos de una ecuación por un escalar no nulo.*
3) *Suma de una ecuación a otra multiplicada por un escalar.*

Cada vez que efectuamos una de esas operaciones en el sistema obtenemos un nuevo sistema con las mismas soluciones. Dos sistemas con las mismas soluciones se llaman *equivalentes.* Efectuando esas operaciones una tras otra de modo sistemático llegamos por fin a un sistema equivalente que puede resolverse a simple vista.

Ilustraremos el método con algunos ejemplos particulares. Se verá entonces cómo se aplica el método en general.

EJEMPLO 1. *Sistema con solución única.* Consideremos el sistema

$$\begin{aligned} 2x - 5y + 4z &= -3 \\ x - 2y + z &= 5 \\ x - 4y + 6z &= 10\,. \end{aligned}$$

Este sistema tiene solución única, $x = 124$, $y = 75$, $z = 31$, que obtendremos por el método de eliminación de Gauss-Jordan. Para evitar trabajo no copiamos las letras x, y, z ni los signos de igualdad, sino que trabajaremos con la *matriz ampliada*

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \tag{2.24}$$

obtenida al adjuntar los segundos miembros del sistema a la matriz de los coeficientes. Las tres operaciones básicas antes mencionadas se efectúan con las filas de la matriz ampliada y se llaman *operaciones fila.* En cualquier fase del proceso podemos poner las letras x, y, z e intercalar los signos de igualdad en las verticales correspondientes obteniendo ecuaciones. Nuestro objetivo es llegar a

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \end{array}\right] \tag{2.25}$$

la matriz ampliada después de una sucesión de operaciones fila. El correspondiente sistema de ecuaciones es $x = 124$, $y = 75$, $z = 31$, que nos da la solución deseada.

El primer paso es obtener un 1 en el vértice superior izquierdo de la matriz. Podemos hacerlo intercambiando la primera fila de la matriz dada (2.24) con la segunda o con la tercera. O bien, podemos multiplicar la primera fila por ½. Intercambiando las filas primera y segunda, obtenemos

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]$$

El paso siguiente consiste en convertir todos los restantes elementos de la primera columna en ceros, dejando el primero intacto. Basta para ello multiplicar la primera fila por -2 y sumar el resultado a la segunda fila. Luego multiplicamos la primera fila por -1 y sumamos el resultado a la tercera. Después de esas dos operaciones, obtenemos

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & -1 & 2 & -13 \\ 0 & -2 & 5 & 5 \end{array}\right]. \tag{2.26}$$

Repetimos ahora el proceso en la matriz reducida $\left[\begin{array}{cc|c} -1 & 2 & -13 \\ -2 & 5 & 5 \end{array}\right]$ que aparece junto a los dos ceros. Podemos obtener 1 en *su* vértice superior izquierdo multiplicando la segunda fila de (2.26) por -1. Esto nos da la matriz

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & -2 & 5 & 5 \end{array}\right].$$

Multiplicando la segunda fila por 2 y sumando el resultado a la tercera, conseguimos

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right]. \tag{2.27}$$

Al llegar aquí, el correspondiente sistema de ecuaciones viene dado por

$$\begin{aligned} x - 2y + z &= 5 \\ y - 2z &= 13 \\ z &= 31 . \end{aligned}$$

Estas ecuaciones pueden resolverse sucesivamente partiendo de la tercera y trabajando hacia atrás, dándonos

$$z = 31,\ \ y = 13 + 2z = 13 + 62 = 75,\ \ \ x = 5 + 2y - z = 5 + 150 - 31 = 124\,.$$

O bien, podemos continuar el proceso de Gauss-Jordan convirtiendo en ceros todos los elementos situados por encima de la diagonal de unos en la segunda y en la tercera columnas. Multiplicando la segunda fila de (2.27) por 2 y sumando el resultado a la primera, obtenemos

$$\left[\begin{array}{ccc|c} 1 & 0 & -3 & 31 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right].$$

Por último, multiplicamos la tercera fila por 3 y sumamos el resultado a la primera fila, y luego multiplicamos la tercera fila por 2 y sumamos el resultado a la segunda con lo que llegamos a la matriz (2.25).

EJEMPLO 2. *Sistema con más de una solución.* Consideremos el siguiente sistema de 3 ecuaciones con 5 incógnitas:

$$\begin{aligned} 2x - 5y + 4z + \ u - v &= -3 \\ x - 2y + \ z - \ u + v &= 5 \\ x - 4y + 6z + 2u - v &= 10\,. \end{aligned} \tag{2.28}$$

La correspondiente matriz ampliada es

$$\left[\begin{array}{ccccc|c} 2 & -5 & 4 & 1 & -1 & -3 \\ 1 & -2 & 1 & -1 & 1 & 5 \\ 1 & -4 & 6 & 2 & -1 & 10 \end{array}\right].$$

Los coeficientes de x, y, z y los segundos miembros son los mismos que los del ejemplo 1. Si efectuamos las mismas operaciones fila que en el ejemplo 1, llegamos a la matriz ampliada

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & -16 & 19 & 124 \\ 0 & 1 & 0 & -9 & 11 & 75 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right].$$

El correspondiente sistema de ecuaciones puede resolverse respecto a x, y, z en función de u y v dándonos

$$\begin{aligned} x &= 124 + 16u - 19v \\ y &= 75 + 9u - 11v \\ z &= 31 + 3u - 4v\,. \end{aligned}$$

Si hacemos $u = t_1$ y $v = t_2$, siendo t_1 y t_2 números reales arbitrarios, y determinamos x, y, z mediante esas ecuaciones, el vector (x, y, z, u, v) de V_5 dado por

$$(x, y, z, u, v) = (124 + 16t_1 - 19t_2\,,\ 75 + 9t_1 - 11t_2\,,\ 31 + 3t_1 - 4t_2\,,\ t_1\,,\ t_2)$$

es una solución. Separando los términos que contienen t_1 y t_2, tendremos:

$$(x, y, z, u, v) = (124, 75, 31, 0, 0) + t_1(16, 9, 3, 1, 0) + t_2(-19, -11, -4, 0, 1)\,.$$

Esta ecuación nos da la solución general del sistema. El vector (124, 75, 31, 0, 0) es una solución particular del sistema no homogéneo (2.28). Los dos vectores (16, 9, 3, 1, 0) y (−19, −11, −4, 0, 1) son soluciones del correspondiente sistema homogéneo. Puesto que son independientes, constituyen una base para el espacio de todas las soluciones del sistema homogéneo.

EJEMPLO 3. *Sistema sin solución.* Consideremos el sistema

(2.29)
$$\begin{aligned} 2x - 5y + 4z &= -3 \\ x - 2y + z &= 5 \\ x - 4y + 5z &= 10\,. \end{aligned}$$

Es idéntico al del ejemplo 1 excepto que en el coeficiente de z en la tercera ecuación ha sido cambiado el 6 por un 5. La matriz ampliada correspondiente es

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 5 & 10 \end{array}\right].$$

Aplicando las mismas operaciones fila usadas en el ejemplo 1 para transformar (2.24) en (2.27), llegamos a la matriz ampliada

(2.30)
$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 0 & 31 \end{array}\right].$$

Cuando la última fila se expresa como ecuación, llegamos a $0 = 31$. Por consiguiente el sistema original no tiene solución puesto que los dos sistemas (2.29) y (2.30) son equivalentes.

En cada uno de los ejemplos anteriores, el número de ecuaciones no excedía al de incógnitas. Si hay más ecuaciones que incógnitas, el proceso de Gauss-Jordan puede aún aplicarse. Por ejemplo, consideremos el sistema del ejemplo 1, que tiene la solución $x = 124$, $y = 75$, $z = 31$. Si adjuntamos una nueva ecuación a este sistema que sea satisfecha por la misma terna, por ejemplo, la ecuación $2x - 3y + z = 54$, entonces el proceso de eliminación nos lleva a la matriz ampliada

$$\left[\begin{array}{ccc|r} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

con una fila de ceros en la parte inferior. Pero si adjuntamos una nueva ecuación que no se satisfaga por la terna (124, 75, 31), por ejemplo la ecuación $x + y + z = 1$, entonces el proceso de eliminación nos conduce a la matriz ampliada de la forma

$$\left[\begin{array}{ccc|r} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \\ 0 & 0 & 0 & a \end{array}\right],$$

donde $a \neq 0$. La última fila nos da una ecuación contradictoria $0 = a$ lo que prueba que el sistema no tiene solución.

2.19 Inversas de matrices cuadradas

Sea $A = (a_{ij})$ una matriz cuadrada $n \times n$, tal que $BA = I$, siendo I la matriz identidad $n \times n$, entonces A se llama *no singular* y B la *inversa* de A por *la izquierda.*

Elegida la base usual de los vectores coordenados unitarios de V_n, sea $T: V_n \to V_n$ la transformación lineal con matriz $m(T) = A$. Tenemos entonces el siguiente

TEOREMA 2.20. *La matriz A es no singular si y sólo si T es invertible. Si $BA = I$, entonces $B = m(T^{-1})$.*

Demostración. Supongamos que A es no singular y que $BA = I$. Demostraremos que $T(x) = O$ implica $x = O$. Dado x tal que $T(x) = O$, sea X la matriz columna $n \times 1$ formada a partir de los componentes de x. Puesto que $T(x) = O$, la matriz producto AX es una matriz columna $n \times 1$ formada por ceros, así que $B(AX)$ es también una matriz columna de ceros. Pero $B(AX) = (BA)X = IX = X$, por lo que todo componente de x es 0. Por consiguiente, T es invertible, y la ecuación $TT^{-1} = I$ implica que $m(T)m(T^{-1}) = I$ o $Am(T^{-1}) = I$. Multiplicando a la izquierda por B, encontramos $m(T^{-1}) = B$. Recíprocamente, si T es invertible entonces $T^{-1}T$ es la transformación idéntica así que $m(T^{-1})m(T)$ es la matriz identidad. Por consiguiente A es no singular y $m(T^{-1})A = I$.

Todas las propiedades de las transformaciones lineales invertibles tienen su contrapartida para las matrices no singulares. En particular, las inversas por la izquierda (si existen) son únicas, y toda inversa por la izquierda es también inversa por la derecha. Dicho de otro modo, si A es no singular y $BA = I$, entonces $AB = I$. Llamamos a B la *inversa* de A y la designamos por A^{-1}. La inversa A^{-1} también es no singular y *su* inversa es A.

Seguidamente demostramos que el problema de la determinación efectiva de los elementos de la inversa de una matriz no singular es equivalente a la resolución de n sistemas lineales no homogéneos.

Sea $A = (a_{ij})$ no singular y sea $A^{-1} = (b_{ij})$ su inversa. Los elementos de A y A^{-1} están ligados por las n^2 ecuaciones.

$$\sum_{k=1}^{n} a_{ik}b_{kj} = \delta_{ij}, \tag{2.31}$$

siendo $\delta_{ij} = 1$ si $i = j$, y $\delta_{ij} = 0$ si $i \neq j$. Para cada valor fijo de j, podemos considerar (2.31) como un sistema no homogéneo de n ecuaciones lineales con n incógnitas $b_{1j}, b_{2j}, \ldots, b_{nj}$. Puesto que A es no singular, cada uno de esos sistemas tiene solución única, la columna j de B. Todos esos sistemas tienen la misma matriz de coeficientes A y difieren tan sólo en sus segundos miembros. Por ejemplo, si A es una matriz 3×3, existen 9 ecuaciones en (2.31) que pueden representarse como 3 sistemas lineales que tienen las siguientes matrices ampliadas:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & 1 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \end{array}\right], \quad \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 1 \\ a_{31} & a_{32} & a_{33} & 0 \end{array}\right], \quad \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 1 \end{array}\right].$$

Si aplicamos el proceso de Gauss-Jordan, llegamos a las respectivas matrices ampliadas

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & b_{11} \\ 0 & 1 & 0 & b_{21} \\ 0 & 0 & 1 & b_{31} \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & b_{12} \\ 0 & 1 & 0 & b_{22} \\ 0 & 0 & 1 & b_{32} \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & b_{13} \\ 0 & 1 & 0 & b_{23} \\ 0 & 0 & 1 & b_{33} \end{array}\right].$$

En la práctica aprovechamos el hecho de que los tres sistemas tienen la misma matriz de coeficientes y resolvemos los tres sistemas de una vez trabajando con la matriz ampliada

$$\left[\begin{array}{ccc|ccc} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array}\right].$$

El proceso de eliminación nos lleva a

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 & b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 & b_{31} & b_{32} & b_{33} \end{array}\right].$$

La matriz de la parte derecha de la barra vertical es la inversa deseada. La de la izquierda es la matriz identidad 3×3.

No es preciso conocer de antemano si A es no singular. Si A es *singular*, podemos aún aplicar el método de Gauss-Jordan, pero ocurre que en el proceso uno de los elementos de la diagonal se convierte en cero, y no será posible transformar A en la matriz identidad.

Un sistema de n ecuaciones lineales con n incógnitas, está representado por

$$\sum_{k=1}^{n} a_{ik} x_k = c_i\,, \qquad i = 1, 2, \ldots, n\,,$$

que se puede escribir, de una manera más simple, como una ecuación matricial

$$AX = C,$$

donde $A = (a_{ij})$ es la matriz de los coeficientes, y X y C matrices columna,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, \qquad C = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix}.$$

Si A es no singular, la única solución del sistema está dada por $X = A^{-1}C$.

2.20 Ejercicios

Aplicando el proceso de Gauss-Jordan a cada uno de los sistemas siguientes, determinar la solución general, si existe.

1. $\begin{aligned} x + y + 3z &= 5 \\ 2x - y + 4z &= 11 \\ -y + z &= 3. \end{aligned}$

2. $\begin{aligned} 3x + 2y + z &= 1 \\ 5x + 3y + 3z &= 2 \\ x + y - z &= 1. \end{aligned}$

3. $\begin{aligned} 3x + 2y + z &= 1 \\ 5x + 3y + 3z &= 2 \\ 7x + 4y + 5z &= 3. \end{aligned}$

4. $\begin{aligned} 3x + 2y + z &= 1 \\ 5x + 3y + 3z &= 2 \\ 7x + 4y + 5z &= 3 \\ x + y - z &= 0. \end{aligned}$

5. $\begin{aligned} 3x - 2y + 5z + u &= 1 \\ x + y - 3z + 2u &= 2 \\ 6x + y - 4z + 3u &= 7. \end{aligned}$

6. $\begin{aligned} x + y - 3z + u &= 5 \\ 2x - y + z - 2u &= 2 \\ 7x + y - 7z + 3u &= 3. \end{aligned}$

7. $\begin{aligned} x + y + 2z + 3u + 4v &= 0 \\ 2x + 2y + 7z + 11u + 14v &= 0 \\ 3x + 3y + 6z + 10u + 15v &= 0. \end{aligned}$

8. $\begin{aligned} x - 2y + z + 2u &= -2 \\ 2x + 3y - z - 5u &= 9 \\ 4x - y + z - u &= 5 \\ 5x - 3y + 2z + u &= 3. \end{aligned}$

9. Demostrar que el sistema $x + y + 2z = 2$, $2x - y + 3z = 2$, $5x - y + az = 6$, tiene solución única si $a \neq 8$. Hallar todas las soluciones cuando $a = 8$.

10. a) Determinar todas las soluciones del sistema

$$\begin{aligned} 5x + 2y - 6z + 2u &= -1 \\ x - y + z - u &= -2. \end{aligned}$$

b) Determinar todas las soluciones del sistema

$$\begin{aligned} 5x + 2y - 6z + 2u &= -1 \\ x - y + z - u &= -2 \\ x + y + z &= 6. \end{aligned}$$

11. Este ejercicio nos indica cómo se determinan todas las matrices no singulares 2×2. Demostrar que

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = (ad - bc)I .$$

Deducir que $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es no singular si y sólo si $ad - bc \neq 0$, en cuyo caso su inversa es

$$\frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Determinar la inversa de cada una de las matrices de los ejercicios del 12 al 16.

12. $\begin{bmatrix} 2 & 3 & 4 \\ 2 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$

13. $\begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{bmatrix}.$

14. $\begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -4 \\ 1 & -4 & 6 \end{bmatrix}.$

15. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$

16. $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{bmatrix}.$

2.21 Ejercicios varios sobre matrices

1. Si una matriz cuadrada tiene una columna de ceros o una fila de ceros, demostrar que es singular.

2. Para cada una de las proposiciones siguientes relativas a matrices $n \times n$, dar una demostración o un contraejemplo.
   a) Si $AB + BA = O$, entonces $A^2B^3 = B^3A^2$.
   b) Si A y B son no singulares, entonces $A + B$ es no singular.
   c) Si A y B son no singulares, entonces AB es no singular.
   d) Si A, B, y $A + B$ son no singulares, entonces $A - B$ es no singular.
   e) Si $A^3 = O$, entonces $A - I$ es no singular.
   f) Si el producto de k matrices $A_1 \ldots A_k$ es no singular, cada una de las matrices A_i es no singular.
3. Si $A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$, hallar una matriz no singular P tal que $P^{-1}AP = \begin{bmatrix} 6 & 0 \\ 0 & -1 \end{bmatrix}$.
4. La matriz $A = \begin{bmatrix} a & i \\ i & b \end{bmatrix}$, donde $i^2 = -1$, $a = \frac{1}{2}(1 + \sqrt{5})$, y $b = \frac{1}{2}(1 - \sqrt{5})$, tiene la propiedad de que $A^2 = A$. Describir en forma completa todas las matrices A, 2×2, con elementos complejos tales que $A^2 = A$.
5. Si $A^2 = A$, demostrar que $(A + I)^k = I + (2^k - 1)A$.
6. La teoría de la relatividad utiliza un conjunto de ecuaciones de la forma $x' = a(x - vt)$, $y' = y$, $z' = z$, $t' = a(t - vx/c^2)$. Aquí v representa la velocidad de un objeto que se mueve, c la velocidad de la luz, y $a = c/\sqrt{c^2 - v^2}$, donde $|v| < c$. La transformación que aplica el vector bi dimensional (x, t) en (x', t') se llama *transformación de Lorentz*. Su matriz relativa a las bases usuales se designa con $L(v)$ y viene dada por

$$L(v) = a\begin{bmatrix} 1 & -v \\ -vc^{-2} & 1 \end{bmatrix}.$$

   Obsérvese que $L(v)$ es no singular y que $L(0) = I$. Demostrar que $L(v)L(u) = L(w)$, siendo $w = (u + v)c^2/(uv + c^2)$. Es decir, el producto de dos transformaciones de Lorentz es otra transformación de Lorentz.
7. Si cambiamos las filas por las columnas en una matriz rectangular A, la nueva matriz así obtenida se llama la *transpuesta* de A y se designa por A^t. Por ejemplo, si tenemos

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \text{entonces } A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

   Demostrar que las transpuestas tienen las propiedades siguientes:
   a) $(A^t)^t = A$. b) $(A + B)^t = A^t + B^t$. c) $(cA)^t = cA^t$.
   d) $(AB)^t = B^tA^t$. e) $(A^t)^{-1} = (A^{-1})^t$ si A es no singular.
8. Una matriz cuadrada A se llama matriz ortogonal si $AA^t = I$. Comprobar que la matriz 2×2 $\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ es ortogonal para cada número real θ. Si A es cualquier matriz ortogonal $n \times n$, demostrar que sus filas, consideradas como vectores de V_n, forman un conjunto ortogonal.
9. Para cada una de las proposiciones siguientes acerca de las matrices $n \times n$, dar una demostración o en su lugar un contraejemplo.
   a) Si A y B son ortogonales, $A + B$ es ortogonal.

b) Si A y B son ortogonales, AB es ortogonal.
c) Si A y AB son ortogonales, B es ortogonal.

10. *Matrices de Hadamard,* llamadas así por Jacques Hadamard (1865-1963), son aquellas matrices $n \times n$ con las propiedades siguientes:
 I. Cada elemento es 1 ó -1.
 II. Cada fila, considerada como un vector de V_n, tiene longitud igual a $\sqrt{n}$.
 III. El producto escalar de dos filas distintas cualesquiera es 0.

Las matrices de Hadamard se presentan en ciertos problemas de geometría y en la teoría de números, y han sido aplicadas recientemente en la codificación óptima para la comunicación espacial. A pesar de su aparente simplicidad, presentan muchos problemas sin resolver. El principal problema no resuelto en este momento es el de determinar todos los valores de n para los que existe una matriz de Hadamard $n \times n$. Este ejercicio da idea de una solución parcial.

a) Determinar todas las matrices de Hadamard 2×2 (hay exactamente 8).
b) Esta parte del ejercicio esboza una demostración sencilla del siguiente teorema: *Si A es una matriz de Hadamard $n \times n$, siendo $n > 2$, entonces n es un múltiplo de* 4. La demostración se basa en dos lemas muy sencillos relativos a los vectores en el espacio de dimensión n. Demostrar cada uno de esos lemas y aplicarlos a las filas de la matriz de Hadamard para demostrar el teorema.

LEMA 1. *Si X, Y, Z son vectores ortogonales de V_n, se tiene*

$$(X + Y) \cdot (X + Z) = \|X\|^2.$$

LEMA 2. *Póngase $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n)$, $Z = (z_1, \ldots, z_n)$. Si cada componente x_i, y_i, z_i es 1 o -1, el producto $(x_i + y_i)(x_i + z_i)$ es 0 ó 4.*

3

DETERMINANTES

3.1 Introducción

En numerosas aplicaciones del Álgebra lineal a la Geometría y al Análisis el concepto de determinante desempeña un papel importante. Este capítulo estudia las propiedades fundamentales de los determinantes y algunas de sus aplicaciones.

En el Volumen I se introdujeron los determinantes de segundo y tercer orden como una notación útil para expresar ciertas fórmulas en forma compacta. Recordemos que el determinante de segundo orden se definió mediante la fórmula

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{3.1}$$

A pesar del parecido de las notaciones, el determinante $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ (escrito con dos barras verticales) es conceptualmente distinto de la matriz $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ (con dos corchetes). El determinante es un *número* asignado a la matriz y que se calcula con la fórmula (3.1). Para acentuar esta conexión también escribimos

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \tag{3.2}$$

Los determinantes de tercer orden se definieron en el Volumen I en función de determinantes de segundo orden mediante la fórmula

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Este capítulo considera el caso más general, el determinante de una matriz cuadrada de orden n para cualquier entero $n \geq 1$. Nuestro punto de vista consiste en tratar el determinante como una función que asigna a cada matriz cuadrada A un número llamado determinante de A y que se indica por $\det A$. Se puede definir esa función por medio de una fórmula que generaliza (3.1) y (3.2). Tal fórmula es una suma que contiene $n!$ productos de elementos de A. Para valores grandes de n esa fórmula es de difícil manejo y se usa poco en la práctica. Es preferible estudiar los determinantes desde otro punto de vista que ponga de manifiesto con mayor claridad sus propiedades esenciales. Tales propiedades, importantes en las aplicaciones, se tomarán como *axiomas* para definir una función determinante. En principio, nuestro plan constará de tres partes: 1) Justificar la elección de los axiomas. 2) Deducir otras propiedades de los determinantes a partir de los axiomas. 3) Demostrar que existe una función y sólo una que satisfaga a tales axiomas.

3.2 Justificación de la elección de los axiomas para una función determinante

En el Volumen I se demostró que el producto mixto de tres vectores A_1, A_2, A_3 en E_3 puede expresarse como el determinante de una matriz cuyas filas son los vectores dados. Así tenemos

$$A_1 \times A_2 \cdot A_3 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

donde $A_1 = (a_{11}, a_{12}, a_{13})$, $A_2 = (a_{21}, a_{22}, a_{23})$, y $A_3 = (a_{31}, a_{32}, a_{33})$.

Si las filas son linealmente independientes el producto mixto no es nulo; el valor absoluto del producto es igual al volumen del paralelepípedo determinado por los tres vectores A_1, A_2, A_3. Si las filas son linealmente dependientes el producto mixto es nulo. En tal caso A_1, A_2, A_3 son coplanarios y el paralelepípedo degenera en una figura plana de volumen nulo.

Algunas de las propiedades del producto mixto justifican la elección de los axiomas para una función determinante en el caso n-dimensional. Para establecer esas propiedades en forma apta para la generalización, consideramos el producto mixto como una función de los tres vectores fila A_1, A_2, A_3. Designamos esta función por d; así pues,

$$d(A_1, A_2, A_3) = A_1 \times A_2 \cdot A_3.$$

Centremos nuestra atención en las propiedades siguientes:

a) *Homogeneidad en cada fila.* Por ejemplo, la homogeneidad en la primera fila establece que

$$d(tA_1, A_2, A_3) = t\, d(A_1, A_2, A_3) \quad \text{para todo escalar } t.$$

b) *Aditividad en cada fila.* Por ejemplo, la aditividad en la segunda fila significa que

$$d(A_1, A_2 + C, A_3) = d(A_1, A_2, A_3) + d(A_1, C, A_3)$$

para todo vector C.

c) *El producto mixto es nulo si dos de las filas son iguales.*

d) *Normalización:*

$$d(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}) = 1, \qquad \text{donde} \quad \boldsymbol{i} = (1, 0, 0), \quad \boldsymbol{j} = (0, 1, 0), \quad \boldsymbol{k} = (0, 0, 1)$$

Cada una de esas propiedades puede comprobarse fácilmente a partir de las propiedades de los productos escalar y vectorial. Algunas son sugeridas por la relación geométrica entre el producto mixto y el volumen del paralelepípedo determinado por los vectores geométricos A_1, A_2, A_3. El significado geométrico de la propiedad aditiva b) en un caso particular es de especial interés. Si tomamos $C = A_1$ en b) el segundo término del segundo miembro es cero en virtud de c), y la relación b) se transforma en

(3.3) $$d(A_1, A_2 + A_1, A_3) = d(A_1, A_2, A_3).$$

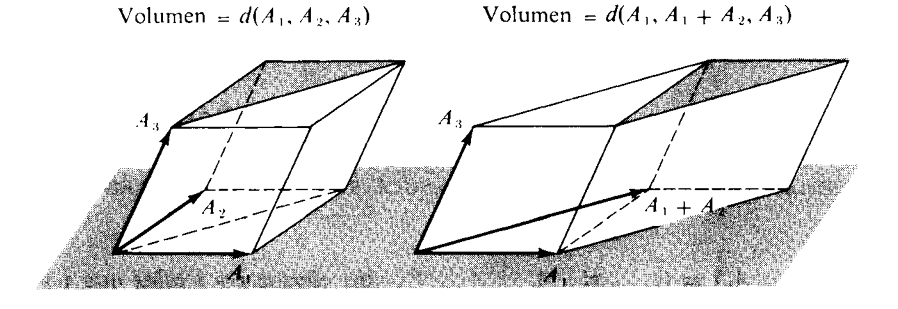


FIGURA 3.1
Interpretación geométrica de la propiedad $d(A_1\ A_2, A_3) = d(A_1, A_1 + A_2, A_3)$. Los dos paralelepípedos tienen volúmenes iguales.

Esta propiedad se pone de manifiesto en la figura 3.1 que representa un paralelepípedo determinado por A_1, A_2, A_3, y otro determinado por $A_1, A_1 + A_2, A_3$. La igualdad (3.3) establece que esos dos paralelepípedos tienen volúmenes iguales. Geométricamente ello es evidente ya que los paralelepípedos tienen alturas iguales y bases de igual área.

3.3 Conjunto de axiomas que definen una función determinante

Las propiedades del producto mixto mencionadas en la sección anterior pueden generalizarse con facilidad y utilizarse como axiomas para determinantes de orden n. Si $A = (a_{ij})$ es una matriz $n \times n$ con elementos reales o complejos, designemos sus filas por $A_1, \ldots, A_n$. Así, la fila i-ésima de A es un vector en E_n dado por

$$A_i = (a_{i1}, a_{i2}, \ldots, a_{in}).$$

Consideremos el determinante como una función de las n filas $A_1, \ldots, A_n$ y designemos su valor por $d(A_1, \ldots, A_n)$ o por $\det A$.

DEFINICIÓN AXIOMÁTICA DE UNA FUNCIÓN DETERMINANTE. *Una función con valores reales o complejos d, definida para cada n componentes ordenadas de los vectores $A_1, \ldots, A_n$ en E_n, se denomina función determinante de orden n si satisface los siguientes axiomas cualesquiera que sean los vectores $A_1, \ldots, A_n$ y C en E_n:*

AXIOMA 1. HOMOGENEIDAD EN CADA FILA. *Si se multiplica la k-ésima fila A_k por un escalar t, el determinante queda también multiplicado por t:*

$$d(\ldots, tA_k, \ldots) = t\, d(\ldots, A_k, \ldots).$$

AXIOMA 2. ADITIVIDAD EN CADA FILA. *Para cada valor de k tenemos*

$$d(A_1, \ldots, A_k + C, \ldots, A_n) = d(A_1, \ldots, A_k, \ldots, A_n) + d(A_1, \ldots, C, \ldots, A_n)$$

AXIOMA 3. EL DETERMINANTE SE ANULA SI DOS FILAS CUALESQUIERA SON IGUALES:

$$d(A_1, \ldots, A_n) = 0 \qquad \text{si} \quad A_i = A_j \qquad (\text{para ciertos } i \text{ y } j \text{ tales que } i \neq j.)$$

AXIOMA 4. EL DETERMINANTE DE LA MATRIZ IDENTIDAD ES IGUAL A 1:

$$d(I_1, \ldots, I_n) = 1, \qquad \text{siendo } I_k \text{ el } k\text{-ésimo vector coordenado unidad}.$$

Los dos primeros axiomas establecen que el determinante de una matriz es una función lineal de cada una de sus filas. Con frecuencia esto se expresa diciendo que el determinante es una función *multilineal* de sus filas. Aplicando reiteradamente la linealidad a la primera fila podemos escribir

$$d\left(\sum_{k=1}^{p} t_k C_k, A_2, \ldots, A_n\right) = \sum_{k=1}^{p} t_k\, d(C_k, A_2, \ldots, A_n),$$

siendo $t_1, \ldots, t_p$ escalares y $C_1, \ldots, C_p$ vectores cualesquiera de E_n.

A veces se utiliza una versión particular del axioma 3:

AXIOMA 3'. EL DETERMINANTE SE ANULA SI DOS FILAS CONSECUTIVAS SON IGUALES:

$$d(A_1, \ldots, A_n) = 0 \qquad \text{si} \quad A_k = A_{k+1} \text{ para un cierto } k = 1, 2, \ldots, n-1.$$

Se demuestra más adelante que para un n dado existe una y sólo una función d que satisface los axiomas 1, 2, 3' y 4. El teorema que sigue da únicamente las propiedades de los determinantes deducidas a partir de los axiomas 1, 2, y 3'. Una de esas propiedades es el axioma 3. Se observará que el axioma 4 no se utiliza en la demostración de este teorema. Tal observación será de utilidad más adelante cuando demostremos la unicidad de la función determinante.

TEOREMA 3.1. *Una función determinante que satisfaga los axiomas* 1, 2, *y* 3' *tiene además las siguientes propiedades:*

a) *El determinante se anula si alguna fila es O:*

$$d(A_1, \ldots, A_n) = 0 \qquad si \quad A_k = O \;\; para\ algún\ k.$$

b) *El determinante cambia de signo si se intercambian dos filas consecutivas:*

$$d(\ldots, A_k, A_{k+1}, \ldots) = -d(\ldots, A_{k+1}, A_k, \ldots).$$

c) *El determinante cambia de signo si dos filas cualesquiera A_i y A_j con $i \neq j$ se intercambian.*

d) *El determinante se anula si son iguales dos filas cualesquiera:*

$$d(A_1, \ldots, A_n) = 0 \qquad si \quad A_i = A_j \qquad para\ un\ par\ i, j,\ con\ i \neq j.$$

e) *El determinante se anula si sus filas son linealmente dependientes.*

Demostración. Para demostrar a) tomemos simplemente $t = 0$ en el axioma 1. Para demostrar b), sea B una matriz que tenga las mismas filas que A excepto las filas k y $k + 1$. Sean las filas B_k y B_{k+1} iguales a $A_k + A_{k+1}$. Entonces según el axioma 3′ $\det B = 0$. Así podemos escribir

$$d(\ldots, A_k + A_{k+1}, A_k + A_{k+1}, \ldots) = 0.$$

Aplicando la propiedad aditiva a las filas k y $k + 1$ podemos escribir esa igualdad en la forma:

$$d(\ldots, A_k, A_k, \ldots) + d(\ldots, A_k, A_{k+1}, \ldots) + d(\ldots, A_{k+1}, A_k, \ldots) + d(\ldots, A_{k+1}, A_{k+1}, \ldots) = 0.$$

Los términos primero y cuarto son nulos en virtud del axioma 3′. Luego el segundo y tercer términos son opuestos, lo que demuestra b).

Para demostrar c) podemos suponer que $i < j$. Se pueden intercambiar las filas A_i y A_j efectuando un número impar de transposiciones de filas *consecutivas*. Primero transponemos la fila A_j sucesivamente con las filas $A_{j-1}, A_{j-2}, \ldots, A_i$. Esto exige $j - i$ transposiciones. Luego la fila A_i la transponemos sucesivamente con las filas $A_{i+1}, A_{i+2}, \ldots, A_{j-1}$. Esto supone $j - i - 1$ transposiciones. Cada transposición o intercambio de filas consecutivas invierte el signo del determinante. El número total de cambios de signo es pues $(j-i)+(j-i-1)=2(j-i)-1$ que es un número *impar,* lo que demuestra c).

Para demostrar d), sea B la matriz obtenida invirtiendo en A las filas A_i y A_j. Puesto que $A_i = A_j$, se tiene $B = A$ y por tanto $\det B = \det A$. Pero, según c) $\det B = -\det A$. Por consiguiente $\det A = 0$.

Para demostrar e) supongamos que existen los escalares $c_1, \ldots, c_n$, no todos nulos, tales que $\sum_{k=1}^{n} c_k A_k = O$. Entonces toda fila A_k con $c_k \neq 0$ puede expresarse como combinación lineal de otras filas. Para mayor sencillez, supongamos que A_1 es combinación lineal de las otras, por ejemplo $A_1 = \sum_{k=2}^{n} t_k A_k$. En virtud de la linealidad de la primera fila tenemos

$$d(A_1, A_2, \ldots, A_n) = d\left(\sum_{k=2}^{n} t_k A_k, A_2, \ldots, A_n\right) = \sum_{k=2}^{n} t_k\, d(A_k, A_2, \ldots, A_n).$$

Pero cada término $d(A_k, A_2, \ldots, A_n)$ de la última suma es cero ya que A_k es igual por lo menos a una de las filas $A_2, \ldots, A_n$. Luego toda la suma es cero. Si la fila A_i es una combinación lineal de las otras filas, razonamos en forma análoga, utilizando la linealidad en la fila i. Esto demuestra e).

3.4 Cálculo de determinantes

Al llegar aquí puede ser instructivo calcular algunos determinantes, utilizando tan sólo los axiomas y las propiedades del teorema 3.1, suponiendo siempre que las funciones determinantes existen. En cada uno de los ejemplos que siguen no aplicamos el axioma 4 hasta el final del cálculo.

EJEMPLO 1. *Determinante de una matriz* 2×2. Queremos demostrar que

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{3.4}$$

Escribamos los vectores fila como combinaciones lineales de los vectores coordenados unitarios $\boldsymbol{i} = (1, 0)$ y $\boldsymbol{j} = (0, 1)$:

$$A_1 = (a_{11}, a_{12}) = a_{11}\boldsymbol{i} + a_{12}\boldsymbol{j}, \qquad A_2 = (a_{21}, a_{22}) = a_{21}\boldsymbol{i} + a_{22}\boldsymbol{j}.$$

Utilizando la linealidad en la primera fila tenemos

$$d(A_1, A_2) = d(a_{11}\boldsymbol{i} + a_{12}\boldsymbol{j}, A_2) = a_{11}\, d(\boldsymbol{i}, A_2) + a_{12}\, d(\boldsymbol{j}, A_2).$$

Utilizándola nuevamente en la segunda fila se obtiene

$$d(\boldsymbol{i}, A_2) = d(\boldsymbol{i}, a_{21}\boldsymbol{i} + a_{22}\boldsymbol{j}) = a_{21}\, d(\boldsymbol{i}, \boldsymbol{i}) + a_{22}\, d(\boldsymbol{i}, \boldsymbol{j}) = a_{22}\, d(\boldsymbol{i}, \boldsymbol{j}),$$

puesto que $d(\boldsymbol{i}, \boldsymbol{i}) = 0$. Del mismo modo encontramos

$$d(\boldsymbol{j}, A_2) = d(\boldsymbol{j}, a_{21}\boldsymbol{i} + a_{22}\boldsymbol{j}) = a_{21}\, d(\boldsymbol{j}, \boldsymbol{i}) = -a_{21}\, d(\boldsymbol{i}, \boldsymbol{j}).$$

Por tanto obtenemos

$$d(A_1, A_2) = (a_{11}a_{22} - a_{12}a_{21})\, d(\boldsymbol{i}, \boldsymbol{j}).$$

Pero $d(\boldsymbol{i}, \boldsymbol{j}) = 1$ según el axioma 4, con lo que $d(A_1, A_2) = a_{11}a_{22} - a_{12}a_{21}$ como se afirmó.

Este razonamiento demuestra que si una función determinante existe para las matrices 2×2, debe tener necesariamente la forma (3.4). Recíprocamente, es fácil comprobar que esa fórmula define, efectivamente, una función determinante de orden 2. Por consiguiente hemos demostrado que existe una y sólo una función determinante de orden 2.

EJEMPLO 2. *Determinante de una matriz diagonal.* Una matriz cuadrada de la forma

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

se denomina *matriz diagonal.* Cada elemento a_{ij} no perteneciente a la diagonal principal, $(i \neq j)$ es cero. Vamos a demostrar que la determinante de A es igual al producto de los elementos pertenecientes a la diagonal

$$\det A = a_{11}a_{22} \cdots a_{nn}. \tag{3.5}$$

La fila k de A es el producto de un escalar por el k-ésimo vector coordenado unidad, $A_k = a_{kk} I_k$.
Aplicando repetidamente la propiedad de homogeneidad llegamos a

$$\det A = d(A_1, \ldots, A_n) = d(a_{11}I_1, \ldots, a_{nn}I_n) = a_{11} \cdots a_{nn}\, d(I_1, \ldots, I_n).$$

Esta fórmula puede escribirse en la forma

$$\det A = a_{11} \cdots a_{nn} \det I,$$

en donde I es la matriz identidad. El axioma 4 nos dice que $\det I = 1$, con lo que obtenemos (3.5).

EJEMPLO 3. *Determinante de una matriz triangular superior.* Una matriz cuadrada de la forma

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

se llama matriz triangular superior. Todos los elementos situados por debajo de

la diagonal principal son nulos. Vamos a demostrar que el determinante de una tal matriz es igual al producto de los elementos de la diagonal principal,

$$\det U = u_{11}u_{22} \cdots u_{nn}.$$

Demostramos primero que $\det U = 0$ si algún elemento u_{ii} de la diagonal principal es $u_{ii} = 0$. Si el último elemento de esa diagonal u_{nn} es cero, la última fila es O y $\det U = 0$ en virtud del teorema 3.1 a). Supongamos, ahora, que sea nulo algún elemento u_{ii} (no precisamente el u_{nn}). Sea por ejemplo $u_{22} = 0$. Entonces cada uno de los $n-1$ vectores fila $U_2, \ldots, U_n$ tienen sus dos primeros componentes nulos. Por tanto esos vectores generan un subespacio cuya dimensión es por lo menos $n-2$. De ahí que esas $n-1$ filas (y por tanto *todas* las filas) sean linealmente dependientes. Según el teorema 3.1 e), $\det U = 0$. Del mismo modo se demuestra que $\det U = 0$, si es cero cualquier elemento de la diagonal principal.

Vamos a demostrar ahora el caso general. Escribamos la primera fila U_1 como una suma de dos vectores-fila,

$$U_1 = V_1 + V_1',$$

siendo $V_1 = [u_{11}, 0, \ldots, 0]$ y $V_1 = [0, u_{12}, \ldots, u_{1n}]$. En virtud de la linealidad en la primera fila tenemos

$$\det U = \det(V_1, U_2, \ldots, U_n) + \det(V_1', U_2, \ldots, U_n).$$

Pero $\det(V_1', U_2, \ldots, U_n) = 0$ ya que éste es el determinante de una matriz triangular superior con un elemento nulo en la diagonal principal. Luego tenemos

$$\det U = \det(V_1, U_2, \ldots, U_n). \tag{3.6}$$

Tratemos ahora el vector-fila U_2 en forma análoga, expresándolo como una suma,

$$U_2 = V_2 + V_2',$$

siendo

$$V_2 = [0, u_{22}, 0, \ldots, 0] \qquad \text{y} \qquad V_2' = [0, 0, u_{23}, \ldots, u_{2n}].$$

Hagamos esto en el segundo miembro de (3.6) y apliquemos la linealidad en la segunda fila con lo que obtenemos

$$\det U = \det(V_1, V_2, U_3, \ldots, U_n), \tag{3.7}$$

ya que $\det(V_1, V_2', U_3, \ldots, U_n) = 0$. Repitiendo el razonamiento para cada una de las filas que siguen en el segundo miembro de (3.7) obtenemos finalmente

$$\det U = \det(V_1, V_2, \ldots, V_n),$$

donde $(V_1, V_2, \ldots, V_n)$ es una matriz diagonal con los mismos elementos que U en la diagonal principal. Por consiguiente, según el ejemplo 2, tenemos

$$\det U = u_{11}u_{22} \cdots u_{nn},$$

como deseábamos demostrar.

EJEMPLO 4. *Cálculo por el método de Gauss-Jordan.* El método de eliminación de Gauss-Jordan para resolver sistemas de ecuaciones lineales es también uno de los mejores métodos para el cálculo de determinantes. Recordemos que el método consiste en la aplicación de tres tipos de operaciones a las filas de la matriz:

1) *Transposición de dos filas.*
2) *Multiplicación de todos los elementos de una fila por un escalar no nulo.*
3) *Adición a una fila del producto de otra por un escalar.*

Efectuando una y otra vez esas operaciones de manera sistemática podemos transformar cualquier matriz cuadrada A en una matriz triangular superior U cuyo determinante sabemos ya calcular. Es fácil determinar la relación entre $\det A$ y $\det U$. Cada vez que se realiza la operación 1) el determinante cambia de signo. Cada vez que se efectúa 2) con un escalar $c \neq 0$, el determinante se multiplica por c. Si se efectúa 3) el determinante no varía. Por consiguiente, si la operación 1) se efectúa p veces y si $c_1, \ldots, c_q$ son los factores escalares que se utilizan en la operación 2), se tiene

$$\det A = (-1)^p (c_1 c_2 \cdots c_q)^{-1} \det U. \tag{3.8}$$

Observemos nuevamente que esta fórmula es consecuencia tan sólo de los tres primeros axiomas. Su demostración no depende del axioma 4.

3.5 El teorema de unicidad

En el ejemplo 3 de la sección anterior se demostró que los axiomas 1, 2, y 3 implican la fórmula $\det U = u_{11}u_{22} \ldots u_{nn} \det I$. Combinando ésta con (3.8) vemos que para toda matriz $A, n \times n$, existe un escalar c (que depende de A) tal que

$$d(A_1, \ldots, A_n) = c\, d(I_1, \ldots, I_n). \tag{3.9}$$

Además, esta fórmula es consecuencia únicamente de los axiomas 1, 2 y 3. A partir de este resultado fácilmente podemos demostrar que no puede existir más de una función determinante.

TEOREMA 3.2. TEOREMA DE UNICIDAD PARA LOS DETERMINANTES. *Sea d una función que satisfaga los cuatro axiomas que caracterizan una función determinante de orden n, y sea f otra función que satisfaga los axiomas* 1, 2 *y* 3. *Entonces, elegido un sistema cualquiera de vectores $A_1, \ldots, A_n$ del espacio n-dimensional tenemos*

(3.10) $$f(A_1, \ldots, A_n) = d(A_1, \ldots, A_n)f(I_1, \ldots, I_n).$$

En particular, si f satisface también el axioma 4 *tenemos $f(A_1, \ldots, A_n) = d(A_1, \ldots, A_n)$.*

Demostración. Sea $g(A_1, \ldots, A_n) = f(A_1, \ldots, A_n) - d(A_1, \ldots, A_n)f(I_1, \ldots, I_n)$. Demostraremos que $g(A_1, \ldots, A_n) = 0$ cualesquiera que sean $A_1, \ldots, A_n$. Puesto que d y f satisfacen ambas los axiomas 1, 2 y 3 lo mismo le ocurre a g. Luego g también satisface la ecuación (3.9) ya que ésa se dedujo a partir de los tres primeros axiomas exclusivamente. Podemos escribir por tanto

(3.11) $$g(A_1, \ldots, A_n) = c\, g(I_1, \ldots, I_n),$$

en donde c es un escalar que depende de A. Tomando $A = I$ en la definición de g y teniendo en cuenta que d satisface el axioma 4 encontramos

$$g(I_1, \ldots, I_n) = f(I_1, \ldots, I_n) - f(I_1, \ldots, I_n) = 0.$$

Por consiguiente, la ecuación (3.11) se convierte en $g(A_1, \ldots, A_n) = 0$. Esto completa la demostración.

3.6 Ejercicios

En este conjunto de ejercicios debe suponerse la existencia de una función determinante. Los determinantes de orden 3 pueden calcularse con la fórmula (3.2).

1. Calcular cada uno de los determinantes siguientes.

$$\text{a)}\ \begin{vmatrix} 2 & 1 & 1 \\ 1 & 4 & -4 \\ 1 & 0 & 2 \end{vmatrix}, \qquad \text{b)}\ \begin{vmatrix} 3 & 0 & 8 \\ 5 & 0 & 7 \\ -1 & 4 & 2 \end{vmatrix}, \qquad \text{c)}\ \begin{vmatrix} a & 1 & 0 \\ 2 & a & 2 \\ 0 & 1 & a \end{vmatrix}.$$

2. Si $\det \begin{bmatrix} x & y & z \\ 3 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} = 1$, calcular el determinante de cada una de las matrices siguientes:

a) $\begin{bmatrix} 2x & 2y & 2z \\ \frac{3}{2} & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, b) $\begin{bmatrix} x & y & z \\ 3x+3 & 3y & 3z+2 \\ x+1 & y+1 & z+1 \end{bmatrix}$, c) $\begin{bmatrix} x-1 & y-1 & z-1 \\ 4 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}$.

3. a) Demostrar que $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)(c-a)(c-b)$.

b) Hallar las correspondientes fórmulas para los determinantes

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix} \quad \text{y} \quad \begin{vmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix}.$$

4. Calcular los determinantes de cada una de las siguientes matrices transformando cada una de ellas en una matriz triangular superior.

a) $\begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$, b) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{bmatrix}$, c) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{bmatrix}$,

d) $\begin{bmatrix} a & 1 & 0 & 0 & 0 \\ 4 & a & 2 & 0 & 0 \\ 0 & 3 & a & 3 & 0 \\ 0 & 0 & 2 & a & 4 \\ 0 & 0 & 0 & 1 & a \end{bmatrix}$, e) $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix}$.

5. Una matriz triangular inferior $A = (a_{ij})$ es una matriz cuadrada que tiene nulos todos los elementos situados por encima de la diagonal principal; esto es, $a_{ij} = 0$ siempre que $i < j$. Demostrar que el determinante de una tal matriz es igual al producto de los elementos de la diagonal principal: $\det A = a_{11}a_{22} \ldots a_{nn}$.

6. Sean f_1, f_2, g_1, g_2 cuatro funciones derivables en un intervalo (a, b). Se define

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ g_1(x) & g_2(x) \end{vmatrix}$$

para cada x de (a, b). Demostrar que

$$F'(x) = \begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & f_2(x) \\ g_1'(x) & g_2'(x) \end{vmatrix}.$$

7. Establecer y demostrar una generalización del ejercicio 6 para el determinante

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ g_1(x) & g_2(x) & g_3(x) \\ h_1(x) & h_2(x) & h_3(x) \end{vmatrix}.$$

8. a) Si $F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix}$, demostrar que $F'(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1''(x) & f_2''(x) \end{vmatrix}$.

b) Establecer y demostrar un resultado análogo para determinantes 3×3, suponiendo la validez de la ecuación (3.2).

9. Sean U y V dos matrices triangulares superiores $n\times n$.

a) Demostrar que $U + V$ y UV son matrices triangulares superiores.

b) Demostrar que $\det(UV) = (\det U)(\det V)$.

c) Si $\det U \neq 0$ demostrar que existe una matriz triangular superior U^{-1} tal que $UU^{-1}=I$, y deducir que $\det(U^{-1}) = 1/\det U$.

10. Calcular $\det A$, $\det(A^{-1})$, y A^{-1} para la siguiente matriz triangular superior:

$$A = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

3.7 Producto de determinantes

En esta sección utilizamos el teorema de unicidad para demostrar que el determinante del producto de dos matrices cuadradas es igual al producto de sus determinantes,

$$\det(AB) = (\det A)(\det B),$$

Recordemos que el producto AB de dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ es la matriz $C = (c_{ij})$ cuyo elemento ij viene dado por la fórmula

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} . \tag{3.12}$$

El producto está definido únicamente si el número de columnas del factor de la izquierda A es igual al número de filas del factor de la derecha B. Esto ocurre si A y B son matrices cuadradas del mismo orden.

En la demostración de la fórmula para el producto se hace uso de una sencilla relación que existe entre las filas de AB y las de A. La establecemos como un lema. Como de costumbre, designamos con A_i la fila i de la matriz A.

LEMA 3.3. *Si A es una matriz $m \times n$ y B una matriz $n \times p$, se tiene*

$$(AB)_i = A_i B .$$

Esto es, la fila i del producto AB es igual al producto de la fila A_i por B.

Demostración. Designemos con B^j la columna j de B y sea $C = AB$. Entonces la suma (3.12) puede considerarse como el producto escalar de la fila i de A por la columna j de B,

$$c_{ij} = A_i \cdot B^j .$$

Por tanto la fila i, C_i, es la matriz fila

$$C_i = [A_i \cdot B^1, A_i \cdot B^2, \ldots, A_i \cdot B^p] .$$

Pero ésta es también el resultado de la multiplicación de la matriz fila A_i por B, ya que

$$A_i B = [a_{i1}, \ldots, a_{in}] \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} = [A_i \cdot B^1, \ldots, A_i \cdot B^p] .$$

Por consiguiente $C_i = A_i B$, lo que demuestra el lema

TEOREMA 3.4. FÓRMULA PARA MULTIPLICAR DETERMINANTES. *Para dos matrices $n \times n$ cualesquiera A y B tenemos*

$$\det(AB) = (\det A)(\det B).$$

Demostración. Puesto que $(AB)_i = A_iB$, tenemos que demostrar que

$$d(A_1B, \ldots, A_nB) = d(A_1, \ldots, A_n)\, d(B_1, \ldots, B_n).$$

Volviendo a utilizar el lema también tenemos $B_i = (IB)_i = I_iB$, donde I es la matriz identidad $n \times n$. Por consiguiente $d(B_1, \ldots, B_n) = d(I_1B, \ldots, I_nB)$, y tenemos que demostrar que

$$d(A_1B, \ldots, A_nB) = d(A_1, \ldots, A_n)\, d(I_1B, \ldots, I_nB).$$

Mantengamos B fija e introduzcamos una función f definida por la fórmula

$$f(A_1, \ldots, A_n) = d(A_1B, \ldots, A_nB).$$

La igualdad que deseamos demostrar establece que

$$f(A_1, \ldots, A_n) = d(A_1, \ldots, A_n) f(I_1, \ldots, I_n). \tag{3.13}$$

Es ahora asunto fácil comprobar que f satisface los axiomas 1, 2 y 3 relativos a una función determinante, con lo que, en virtud del teorema de unicidad, la igualdad (3.13) es válida para cualquier matriz A. Esto completa la demostración.

En las dos seccciones que siguen se exponen aplicaciones de la fórmula del producto.

3.8 Determinante de la matriz inversa de una matriz no singular

Recordemos que una matriz cuadrada A se llama no singular si tiene una inversa B por la izquierda tal que $BA = I$. Si existe una inversa por la izquierda es única y es también inversa por la derecha, $AB = I$. Designemos la inversa por A^{-1}. La relación entre $\det A$ y $\det A^{-1}$ es la que podía esperarse.

TEOREMA 3.5. *Si una matriz A es no singular, entonces $\det A \neq 0$ y tenemos*

$$\det A^{-1} = \frac{1}{\det A}. \tag{3.14}$$

Demostración. Según la fórmula del producto tenemos

$$(\det A)(\det A^{-1}) = \det (AA^{-1}) = \det I = 1 .$$

Luego $\det A \neq 0$ y (3.14) es cierta.

El teorema 3.5 demuestra que la no anulación de $\det A$ es una condición necesaria para que A sea no singular. Más adelante demostraremos que esa condición es también suficiente. Esto es, si $\det A \neq 0$ existe A^{-1}.

3.9 Determinantes e independencia de vectores

Del teorema 3.5 puede deducirse un sencillo criterio para averiguar la independencia de vectores.

TEOREMA 3.6. *Un conjunto de n vectores $A_1, \ldots, A_n$ de un espacio n-dimensional es linealmente independiente si y sólo si $d(A_1, \ldots, A_n) \neq 0$.*

Demostración. Ya demostramos en el teorema 3.2 e) que la dependencia lineal implica $d(A_1, \ldots, A_n) = 0$. Para demostrar el recíproco, se supone que $A_1, \ldots, A_n$ son linealmente independientes y se prueba que $d(A_1, \ldots, A_n) \neq 0$.

Designemos con V_n el espacio lineal de las n-plas de escalares. Puesto que $A_1, \ldots, A_n$ son n elementos independientes en un espacio n-dimensional, formarán una base para V_n. Según el teorema 2.12 existe una transformación lineal $T: V_n \to V_n$ que aplica esos n vectores en los vectores coordenadas unidad,

$$T(A_k) = I_k \quad \text{para} \quad k = 1, 2, \ldots, n .$$

Existe pues una matriz B, $n \times n$, tal que

$$A_k B = I_k \quad \text{para} \quad k = 1, 2, \ldots, n .$$

Pero según el lema 3.3 tenemos $A_k B = (AB)_k$, siendo A la matriz con las filas $A_1, \ldots, A_n$. Luego $AB = I$, así que A es no singular y $\det A \neq 0$.

3.10 Determinante de una matriz diagonal en bloques

Una matriz cuadrada C de la forma

$$C = \begin{bmatrix} A & O \\ O & B \end{bmatrix},$$

en la que A y B son matrices cuadradas y cada O representa una matriz de ceros, se denomina *matriz diagonal en bloques* con dos bloques A y B. Como ejemplo sea la matriz 5×5.

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 4 & 5 & 6 \\ 0 & 0 & 7 & 8 & 9 \end{bmatrix}.$$

Los bloques diagonales son en este caso

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

El teorema que sigue demuestra que el determinante de una matriz diagonal en bloques es igual al producto de los determinantes de sus bloques diagonales.

TEOREMA 3.7. *Para dos matrices cuadradas cualesquiera A y B se tiene*

$$\det \begin{bmatrix} A & O \\ O & B \end{bmatrix} = (\det A)(\det B). \tag{3.15}$$

Demostración. Supongamos que A es $n \times n$ y B es $m \times m$. Observemos que la matriz diagonal en bloques dada, puede expresarse como un producto de la forma

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & I_m \end{bmatrix} \begin{bmatrix} I_n & O \\ O & B \end{bmatrix}$$

donde I_n e I_m son matrices identidad de órdenes n y m respectivamente. Por consiguiente, según la fórmula del producto para determinantes tenemos

$$\det \begin{bmatrix} A & O \\ O & B \end{bmatrix} = \det \begin{bmatrix} A & O \\ O & I_m \end{bmatrix} \det \begin{bmatrix} I_n & O \\ O & B \end{bmatrix}. \tag{3.16}$$

Consideremos ahora el determinante $\det \begin{bmatrix} A & O \\ O & I_m \end{bmatrix}$ como una función de n filas de A. Esto es posible debido al bloque de ceros del ángulo superior derecho. Resulta fácil comprobar que esta función satisface los cuatro axiomas de una función determinante de orden n. Por consiguiente, en virtud del teorema de unicidad, debe ser

$$\det \begin{bmatrix} A & O \\ O & I_m \end{bmatrix} = \det A .$$

Análogamente se demuestra que $\det \begin{bmatrix} I_n & O \\ O & B \end{bmatrix} = \det B$. Luego (3.16) implica (3.15).

3.11 Ejercicios

1. Para cada una de las proposiciones siguientes relativas a matrices cuadradas, dar una demostración o poner un contraejemplo.

   a) $\det (A + B) = \det A + \det B$.
   b) $\det \{(A + B)^2\} = \{\det (A + B)\}^2$
   c) $\det \{(A + B)^2\} = \det (A^2 + 2AB + B^2)$
   d) $\det \{(A + B)^2\} = \det (A^2 + B^2)$.

2. a) Extender el teorema 3.7 a matrices con tres bloques diagonales:

$$\det \begin{bmatrix} A & O & O \\ O & B & O \\ O & O & C \end{bmatrix} = (\det A)(\det B)(\det C) .$$

   b) Establecer y demostrar una generalización del teorema para matrices con un número cualquiera de bloques diagonales.

3. Sea $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & b & c & d \\ e & f & g & h \end{bmatrix}$, $B = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Demostrar que $\det A = \det \begin{bmatrix} c & d \\ g & h \end{bmatrix}$ y que $\det B = \det \begin{bmatrix} a & b \\ e & f \end{bmatrix}$.

4. Establecer y demostrar una generalización del ejercicio 3 para matrices $n \times n$.

5. Sea $A = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ e & f & g & h \\ x & y & z & w \end{bmatrix}$. Demostrar que $\det A = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \det \begin{bmatrix} g & h \\ z & w \end{bmatrix}$.

6. Establecer y demostrar una generalización del ejercicio 5 para matrices $n \times n$ de la forma

$$A = \begin{bmatrix} B & O \\ C & D \end{bmatrix}$$

donde B, C, D son matrices cuadradas y O es una matriz de ceros.

7. Aplicar el teorema 3.6 para determinar si los siguientes conjuntos de vectores son linealmente dependientes o independientes.

a) $A_1 = (1, -1, 0)$, $A_2 = (0, 1, -1)$, $A_3 = (2, 3, -1)$.
b) $A_1 = (1, -1, 2, 1)$, $A_2 = (-1, 2, -1, 0)$, $A_3 = (3, -1, 1, 0)$, $A_4 = (1, 0, 0, 1)$.
c) $A_1 = (1, 0, 0, 0, 1)$, $A_2 = (1, 1, 0, 0, 0)$, $A_3 = (1, 0, 1, 0, 1)$, $A_4 = (1, 1, 0, 1, 1)$, $A_5 = (0, 1, 0, 1, 0)$.

3.12 Fórmulas para desarrollar determinantes. Menores y cofactores

No hemos demostrado todavía la existencia de una función determinante, excepto en el caso 2×2. En esta sección utilizamos la linealidad y el teorema de unicidad para demostrar que, si existen determinantes, pueden calcularse mediante una fórmula que expresa cualquier determinante de orden n como una combinación lineal de determinantes de orden $n-1$. La igualdad (3.2) de la sección 3.1 es un ejemplo de esa fórmula en el caso 3×3. La fórmula general sugerirá un método para demostrar por inducción la existencia de funciones determinantes.

Toda fila de una matriz A, $n \times n$, puede expresarse como una combinación lineal de los n vectores coordenados unidad $I_1, \ldots, I_n$. Por ejemplo, la primera fila A_1 puede escribirse como sigue:

$$A_1 = \sum_{j=1}^{n} a_{1j} I_j .$$

Puesto que los determinantes son lineales respecto a la primera fila tenemos

$$(3.17) \quad d(A_1, A_2, \ldots, A_n) = d\left(\sum_{j=1}^{n} a_{1j} I_j, A_2, \ldots, A_n\right) = \sum_{j=1}^{n} a_{1j}\, d(I_j, A_2, \ldots, A_n).$$

Por consiguiente para calcular $\det A$ basta calcular $d(I_j, A_2, \ldots, A_n)$ para cada vector coordenado unidad I_j.

Utilicemos la notación A'_{1j} para representar la matriz obtenida de A reemplazando la primera fila A_1 por el vector unitario I_j. Por ejemplo, si $n = 3$ existen tres matrices de ese tipo:

$$A'_{11} = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad A'_{12} = \begin{bmatrix} 0 & 1 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad A'_{13} = \begin{bmatrix} 0 & 0 & 1 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Observemos que $\det A'_{1j} = d(I_j, A_2, \ldots, A_n)$. La igualdad (3.17) puede ahora escribirse en la forma

$$(3.18) \qquad \det A = \sum_{j=1}^{n} a_{1j} \det A'_{1j}.$$

Ésta es una *fórmula de desarrollo;* y en ella se expresa el determinante de A como una combinación lineal de los elementos de su primera fila.

El razonamiento utilizado para deducir (3.18) puede aplicarse a la fila k en lugar de hacerlo con la primera. El resultado es una fórmula de desarrollo en función de los elementos de la fila k.

TEOREMA 3.8. DESARROLLO POR COFACTORES. *Sea A'_{kj} la matriz obtenida de la matriz A al sustituir la fila k de A_k por el vector coordenado unidad I_j. Se obtiene entonces la fórmula de desarrollo*

$$(3.19) \qquad \det A = \sum_{j=1}^{n} a_{kj} \det A'_{kj}$$

que expresa el determinante de A como combinación lineal de los elementos de la fila k. El número $\det A'_{kj}$ *se llama cofactor o adjunto del elemento a_{kj}.*

En el teorema que sigue demostraremos que cada cofactor es, salvo el signo, igual al determinante de una matriz de orden $n-1$. Esas matrices se llaman *menores.*

DEFINICIÓN. *Dada una matriz cuadrada A de orden $n \geq 2$, la matriz cuadrada de orden $n-1$ obtenida suprimiendo la fila k y la columna j de A se llama menor k,j de A y se designa por A_{kj}.*

EJEMPLO. Una matriz $A = (a_{kj})$ de orden 3 tiene nueve menores. Tres de ellas son

$$A_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad A_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad A_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

La igualdad (3.2) expresa el determinante de una matriz 3×3 como una combinación lineal de los determinantes de esas tres menores. La fórmula puede escribirse como sigue:

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}.$$

El siguiente teorema extiende esa fórmula al caso $n \times n$ para cualquier $n \geq 2$.

TEOREMA 3.9. DESARROLLO POR LAS MENORES CORRESPONDIENTES A LA FILA k. *Para cualquier matriz A, $n \times n$, $n \geq 2$, el cofactor de a_{kj} está relacionado con la menor A_{kj} por la fórmula*

$$\det A'_{kj} = (-1)^{k+j} \det A_{kj}. \tag{3.20}$$

Por consiguiente, el desarrollo del $\det A$ *en función de los elementos de la fila k viene dado por la fórmula*

$$\det A = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj}. \tag{3.21}$$

Demostración. Consideremos primero el caso particular $k = j = 1$. La matriz A'_{11} tiene la forma

$$A'_{11} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}.$$

Aplicando las operaciones elementales del tipo (3) a las filas, podemos convertir en ceros todos los elementos situados debajo del 1 en la primera columna, dejando invariables los demás elementos. Por ejemplo, si multiplicamos la primera fila de A'_{11} por $-a_{21}$ y sumamos el resultado a la segunda fila, la nueva segunda fila será $(0, a_{22}, a_{23}, \ldots, a_{2n})$. Por medio de una sucesión de esas operaciones elementales

con las filas obtenemos una nueva matriz que designaremos con A_{11}^0 y que tiene la forma

$$A_{11}^0 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ 0 & a_{32} & \cdots & a_{3n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Puesto que las operaciones con las filas del tipo (3) no modifican el determinante tenemos

$$\det A_{11}^0 = \det A'_{11}. \tag{3.22}$$

Pero A_{11}^0 es un matriz diagonal en bloques, así que, en virtud del teorema 3.7, tenemos

$$\det A_{11}^0 = \det A_{11},$$

siendo A_{11} la menor 1,1 de A,

$$A_{11} = \begin{bmatrix} a_{22} & \cdots & a_{2n} \\ a_{32} & \cdots & a_{3n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Por consiguiente $\det A'_{11} = \det A_{11}$, lo que demuestra (3.20) para $k = j = 1$.

Consideremos seguidamente el caso particular $k = 1$, j cualquiera, y demostremos que

$$\det A'_{1j} = (-1)^{j-1} \det A_{1j}. \tag{3.23}$$

Una vez demostrada (3.23) la fórmula (3.20) más general resulta inmediatamente, debido a que la matriz A'_{kj} puede transformarse en una matriz de la forma B'_{1j} mediante $k-1$ intercambios sucesivos de filas adyacentes. El determinante cambia de signo en cada transposición, así que

$$\det A'_{kj} = (-1)^{k-1} \det B'_{1j}, \tag{3.24}$$

siendo B una matriz $n \times n$ cuya primera fila es I_j y cuya menor $1,j$, B_{1j}, es A_{kj}. Según (3.23), tenemos

$$\det B'_{1j} = (-1)^{j-1} \det B_{1j} = (-1)^{j-1} \det A_{kj},$$

con lo que (3.24) nos da

$$\det A'_{kj} = (-1)^{k-1}(-1)^{j-1} \det A_{kj} = (-1)^{k+j} \det A_{kj}.$$

Por lo tanto si demostramos (3.23) también demostramos (3.20).

Vamos ahora a demostrar (3.23). La matriz A'_{1j} tiene la forma

$$A'_{1j} = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}.$$

Mediante operaciones elementales con las filas del tipo (3) introduzcamos una columna de ceros debajo del 1 y la transformamos en

$$A^0_{1j} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{21} & \cdots & a_{2,j-1} & 0 & a_{2,j+1} & \cdots & a_{2n} \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix}.$$

Como antes, el determinante no varía, así que $\det A^0_{1j} = \det A'_{1j}$. La matriz menor A_{1j} tiene la forma

$$A_{1j} = \begin{bmatrix} a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix}.$$

Consideremos seguidamente el $\det A^0_{1j}$ como una función de las $n-1$ filas de A_{1j}, pongamos $\det A^0_{1j} = f(A_{1j})$. La función f satisface los tres primeros axiomas de

una función determinante de orden $n-1$. Por consiguiente, según el teorema de unicidad podemos escribir

$$f(A_{1j}) = f(J) \det A_{1j}\,, \tag{3.25}$$

en donde J es la matriz identidad de orden $n-1$. En consecuencia, para demostrar (3.23) debemos demostrar que $f(J) = (-1)^{j-1}$. Por definición $f(J)$ es el determinante de la matriz

$$C = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & & 0 & 0 & 0 & & 0 \\ \cdot & \ddots & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \cdot & & \cdot & \cdot & \cdot & \ddots & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \leftarrow \text{Fila } j$$

$\uparrow$ Columna j

Los elementos situados en las líneas oblicuas son todos iguales a 1. Los restantes elementos no escritos son todos 0. Intercambiando la primera fila de C sucesivamente con las filas $2, 3, \ldots, j$ llegamos a la matriz identidad I de orden $n \times n$ después de $j-1$ transposiciones. El determinante cambia de signo en cada transposición, así que $\det C = (-1)^{j-1}$. Luego $f(J) = (-1)^{j-1}$, lo que demuestra (3.23) y por tanto (3.20).

3.13 Existencia de la función determinante

En esta sección utilizamos la inducción respecto de n, que es el orden de la matriz, para demostrar la existencia de las funciones determinantes de cualquier orden. Para $n=2$ ya hemos demostrado que existe una función determinante. Para el caso $n=1$ aceptamos por definición que $\det [a_{11}] = a_{11}$.

Suponiendo que existe una función determinante de orden $n-1$, parece lógico que una función determinante de orden n fuera una de las fórmulas de desarrollo del teorema 3.9, por ejemplo, el desarrollo mediante las menores de la primera fila. No obstante, resulta más sencillo comprobar los axiomas si utilizamos una fórmula análoga expresada en función de las menores de la primera columna.

TEOREMA 3.10. *Supongamos la existencia de determinantes de orden $n-1$. Para una matriz $n\times n$ cualquiera $A=(a_{jk})$, sea f la función definida por la*

$$(3.26)\qquad f(A_1, \ldots, A_n) = \sum_{j=1}^{n}(-1)^{j+1}a_{j1}\det A_{j1}.$$

Entonces f satisface los cuatro axiomas de una función determinante de orden n. Por consiguiente, por inducción, los determinantes de orden n existen para todo n.

Demostración. Consideremos cada término de la suma (3.26) como una función de las filas de A y escribamos

$$f_j(A_1, \ldots, A_n) = (-1)^{j+1}a_{j1}\det A_{j1}.$$

Si comprobamos que cada f_j satisface los axiomas 1 y 2 también f los satisfará.

Multipliquemos la primera fila de A por un escalar t. Tal multiplicación no afecta a la menor A_{11} puesto que ésta no incluye la primera fila. El coeficiente a_{11} está multiplicado por t, así que tenemos

$$f_1(tA_1, A_2, \ldots, A_n) = ta_{11}\det A_{11} = tf_1(A_1, \ldots, A_n).$$

Si $j > 1$ la primera fila de cada menor A_{j1} queda multiplicada por t y el coeficiente a_{j1} no queda afectado, por lo que otra vez tenemos

$$f_j(tA_1, A_2, \ldots, A_n) = tf_j(A_1, A_2, \ldots, A_n).$$

Por lo tanto, cada f_j es homogénea en la primera fila.

Si se multiplica por t la fila k de A, siendo $k > 1$, la menor A_{k1} no se modifica pero a_{k1} queda multiplicado por t, por lo que f_k es homogénea en la fila k. Si $j \neq k$, el coeficiente a_{j1} no se afecta pero alguna fila de A_{j1} queda multiplicada por t. Luego toda f_j es homogénea en la fila k.

Un razonamiento parecido demuestra que cada f_j es aditiva respecto a cualquier fila, por lo cual f satisface los axiomas 1 y 2. Seguidamente demostramos que f satisface el axioma 3', la versión particular del axioma 3. Del teorema 3.1, se deduce entonces que f satisface el axioma 3.

Para comprobar que f satisface el axioma 3', supongamos que sean iguales dos filas consecutivas de A, sean $A_k = A_{k+1}$. Entonces, excepto las menores A_{k1} y $A_{k+1,1}$, cada menor A_{j1} tiene dos filas iguales de modo que $\det A_{j1} = 0$. Por lo tanto la suma (3.26) consta tan sólo de los dos términos correspondientes a $j = k$ y $j = k + 1$,

$$(3.27)\qquad f(A_1, \ldots, A_n) = (-1)^{k+1}a_{k1}\det A_{k1} + (-1)^{k+2}a_{k+1,1}\det A_{k+1,1}.$$

Pero $A_{k1} = A_{k+1,1}$ y $a_{k1} = a_{k+1,1}$ ya que $A_k = A_{k+1}$. Por consiguiente, los dos términos de (3.27) difieren tan sólo en el signo, por lo que $f(A_1, \ldots, A_n) = 0$. Así pues, f satisface el axioma 3'.

Por último, comprobamos que f satisface el axioma 4. Cuando $A = I$ tenemos $a_{11} = 1$ y $a_{j1} = 0$ para $j > 1$. Asimismo, A_{11} es la matriz identidad de orden $n-1$, por lo cual cada término de (3.26) es cero salvo el primero, que es igual a 1. Luego $f(I_1, \ldots, I_n) = 1$ por lo que f satisface el axioma 4.

En la demostración anterior podría haberse muy bien utilizado una función f definida en función de las menores A_{jk} de la columna k en lugar de las menores A_{j1} de la primera columna. En efecto, si ponemos

$$f(A_1, \ldots, A_n) = \sum_{j=1}^{n} (-1)^{j+k} a_{jk} \det A_{jk}, \tag{3.28}$$

exactamente el mismo tipo de demostración hace ver que esa f satisface los cuatro axiomas de una función determinante. Puesto que las funciones determinantes son únicas, las fórmulas de desarrollo (3.28) y las de (3.21) son todas iguales a $\det A$.

Las fórmulas (3.28) no sólo establecen la existencia de las funciones determinantes sino que también ponen de manifiesto un nuevo aspecto de la teoría de los determinantes — una conexión entre las propiedades relativas a filas y las propiedades relativas a columnas. Tal conexión se expone en la sección siguiente.

3.14 Determinante de una matriz transpuesta

Asociada a cada matriz A se considera otra matriz llamada *transpuesta* de A y designada por A^t. Las filas de A^t son las columnas de A. Por ejemplo, si

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \text{entonces } A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Una definición formal puede darse así:

DEFINICIÓN DE TRANSPUESTA. *La transpuesta de una matriz,* $m \times n$, $A = (a_{ij})_{i,j=1}^{m,n}$ *es la matriz,* $n \times m$, A^t *cuyo elemento* i,j *es* a_{ji}.

Si bien la transposición puede aplicarse a cualquier matriz rectangular nos referiremos en primer lugar a las matrices cuadradas. A continuación demostramos que la transposición de una matriz cuadrada no modifica su determinante.

TEOREMA 3.11. *Para toda matriz A, $n \times n$, tenemos*

$$\det A = \det A^t.$$

Demostración. Se demuestra por inducción en n. Para $n = 1$ y $n = 2$ el resultado se comprueba inmediatamente. Supongamos, entonces, que el teorema es cierto para matrices de orden $n-1$. Sean $A = (a_{ij})$ y $B_{1j} = A^t = (b_{ij})$. Desarrollando $\det A$ por sus menores correspondientes a la primera columna y $\det B$ por sus menores relativos a la primera fila tenemos

$$\det A = \sum_{j=1}^{n} (-1)^{j+1} a_{j1} \det A_{j1}, \qquad \det B = \sum_{j=1}^{n} (-1)^{j+1} b_{1j} \det B_{1j}.$$

Pero por la definición de transpuesta tenemos $b_{1j} = a_{j1}$ y $B_{1j} = (A_{j1})^t$. Puesto que hemos supuesto que el teorema es cierto para matrices de orden $n-1$ tenemos $\det B_{ij} = \det A_{j1}$. Luego las sumas anteriores coinciden término a término, con lo que $\det A = \det B$.

3.15 La matriz cofactor

El teorema 3.5 puso de manifiesto que si A es no singular, $\det A \neq 0$. El teorema siguiente demuestra el recíproco. Esto es, si $\det A \neq 0$, existe A^{-1}. Además, da una fórmula para expresar A^{-1} en función de una matriz formada con los cofactores de los elementos de A.

En el teorema 3.9 se demostró que el cofactor de a_{ij} es igual a $(-1)^{i+j} \det A_{ij}$, siendo A_{ij} la menor i,j de A. Designemos ese cofactor por $\operatorname{cof} a_{ij}$. Así pues, por definición,

$$\operatorname{cof} a_{ij} = (-1)^{i+j} \det A_{ij}.$$

DEFINICION DE MATRIZ COFACTOR. *La matriz cuyo elemento i,j es* $\operatorname{cof} a_{ij}$ *se llama la matriz cofactor de A (*) y se designa por* $\operatorname{cof} A$. *Así pues, tenemos*

$$\operatorname{cof} A = (\operatorname{cof} a_{ij})_{i,j=1}^{n} = ((-1)^{i+j} \det A_{ij})_{i,j=1}^{n}.$$

El teorema que sigue demuestra que el producto de A por la transpuesta de su matriz cofactor es, salvo un factor escalar, la matriz identidad I.

(*) En la literatura sobre matrices, especialmente en los tratados clásicos, la transpuesta de la matriz cofactor se llama *adjunta* de A. No obstante, la nomenclatura actual reserva la palabra *adjunta* para otra cosa completamente distinta, que se discutirá en la sección 5.8.

TEOREMA 3.12. *Para cualquier matriz A, $n \times n$, con $n \geq 2$ tenemos*

$$A(\operatorname{cof} A)^t = (\det A)I. \tag{3.29}$$

En particular, si $\det A \neq 0$ *existe la inversa de A y viene dada por*

$$A^{-1} = \frac{1}{\det A}(\operatorname{cof} A)^t.$$

Demostración. Utilizando el teorema 3.9 expresamos $\det A$ en función de sus cofactores relativos a la fila k por la fórmula

$$\det A = \sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{kj}. \tag{3.30}$$

Mantengamos fija k y apliquemos esta relación a una nueva matriz B cuya fila i es igual a la fila k de A para alguna $i \neq k$, y cuyas filas restantes son las mismas que las de A. Entonces $\det B = 0$ debido a que las filas i y k de B son iguales. Expresando $\det B$ en función de sus cofactores relativos a la fila i tenemos

$$\det B = \sum_{j=1}^{n} b_{ij} \operatorname{cof} b_{ij} = 0. \tag{3.31}$$

Pero ya que la fila i de B es igual a la fila k de A tenemos

$$b_{ij} = a_{kj} \qquad \text{y} \qquad \operatorname{cof} b_{ij} = \operatorname{cof} a_{ij} \quad \text{para cualquier } j.$$

Luego (3.31) establece que

$$\sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{ij} = 0 \qquad \text{si} \quad k \neq i. \tag{3.32}$$

Las igualdades (3.30) y (3.32) pueden escribirse conjuntamente así:

$$\sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{ij} = \begin{cases} \det A & \text{si} \quad i = k \\ 0 & \text{si} \quad i \neq k. \end{cases} \tag{3.33}$$

Pero la suma que aparece en el primer miembro de (3.33) es el elemento k,i del producto $A(\operatorname{cof} A)^t$. Por consiguiente (3.33) implica (3.29).

Como corolario directo de los teoremas 3.5 y 3.12 tenemos la siguiente condición necesaria y suficiente para que una matriz cuadrada sea no singular.

TEOREMA 3.13. *Una matriz cuadrada A es no singular si y sólo si $\det A \neq 0$.*

3.16 Regla de Cramer

El teorema 3.12 puede también utilizarse para resolver un sistema de ecuaciones lineales con matriz no singular. Las fórmulas que se obtienen constituyen la llamada *regla de Cramer,* en memoria del matemático suizo Gabriel Cramer (1704-1752).

TEOREMA 3.14. REGLA DE CRAMER. *Si un sistema de n ecuaciones lineales con n incógnitas $x_1, \ldots, x_n$,*

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \qquad (i = 1, 2, \ldots, n) \tag{3.34}$$

tiene la matriz de los coeficientes $A = (a_{ij})$ no singular, existe una solución única para el sistema dada por las fórmulas

$$x_j = \frac{1}{\det A} \sum_{k=1}^{n} b_k \operatorname{cof} a_{kj}, \quad \textit{para} \quad j = 1, 2, \ldots, n.$$

Demostración. El sistema se puede poner en forma matricial

$$AX = B,$$

en donde X y B son matrices columna, $X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$, $\quad B = \begin{bmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix}$. Puesto que A es no singular existe una solución única X dada por

$$X = A^{-1}B = \frac{1}{\det A} (\operatorname{cof} A)^t B. \tag{3.35}$$

Las fórmulas (3.34) se obtienen igualando los componentes en (3.35).

Debe observarse que la fórmula para x_j en (3.34) se puede expresar como cociente de dos determinantes,

$$x_j = \frac{\det C_j}{\det A},$$

en donde C_j es la matriz obtenida de A al reemplazar la columna j de A por la matriz columna B.

3.17 Ejercicios

1. Determinar la matriz cofactor de cada una de las matrices siguientes:

$$\text{a)}\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad \text{b)}\ \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 1 \\ -1 & -2 & 0 \end{bmatrix}, \qquad \text{c)}\ \begin{bmatrix} 3 & 1 & 2 & 4 \\ 2 & 0 & 5 & 1 \\ 1 & -1 & -2 & 6 \\ -2 & 3 & 2 & 3 \end{bmatrix}.$$

2. Determinar la inversa de cada una de las matrices no singulares del ejercicio 1.
3. Hallar todos los valores del escalar λ para los que la matriz $\lambda I - A$ es singular, si A es igual a

$$\text{a)}\ \begin{bmatrix} 0 & 3 \\ 2 & -1 \end{bmatrix}, \qquad \text{b)}\ \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{bmatrix}, \qquad \text{c)}\ \begin{bmatrix} 11 & -2 & 8 \\ 19 & -3 & 14 \\ -8 & 2 & -5 \end{bmatrix}.$$

4. Si A es una matriz $n \times n$ con $n \geq 2$, demostrar cada una de las propiedades de su matriz cofactor:
   a) $\operatorname{cof}(A^t) = (\operatorname{cof} A)^t$. b) $(\operatorname{cof} A)^t A = (\det A)I$.
   c) $A(\operatorname{cof} A)^t = (\operatorname{cof} A)^t A$ (A es permutable con la transpuesta de su matriz cofactor).
5. Resolver mediante la regla de Cramer cada uno de los sistemas que siguen:

a) $x + 2y + 3z = 8, \quad 2x - y + 4z = 7, \quad -y + z = 1.$
b) $x + y + 2z = 0, \quad 3x - y - z = 3, \quad 2x + 5y + 3z = 4.$

6. a) Justificar que cada una de las ecuaciones que siguen es la ecuación cartesiana de una recta, en el plano xy, que pasa por los puntos (x_1, y_1) y (x_2, y_2).

$$\det \begin{bmatrix} x - x_1 & y - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{bmatrix} = 0; \qquad \det \begin{bmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} = 0.$$

   b) Establecer y demostrar relaciones análogas, en el espacio de tres dimensiones, para un plano que pasa por tres puntos.
   c) Establecer y demostrar relaciones análogas en el plano xy para una circunferencia que pasa por tres puntos no alineados.

7. Dadas n^2 funciones f_{ij} cada una derivable en un intervalo (a, b), definir $F(x) = \det [f_{ij}(x)]$ para cada x en (a, b). Demostrar que la derivada $F'(x)$ es una suma de n determinantes,

$$F'(x) = \sum_{i=1}^{n} \det A_i(x),$$

donde $A_i(x)$ es la matriz que se obtiene derivando las funciones de la fila i de $[f_{ij}(x)]$.

8. Una matriz $n \times n$ de funciones de la forma $W(x) = [u_j^{(i-1)}(x)]$, en la que cada fila después de la primera es la derivada de la fila anterior, se llama *matriz wronskiana* en recuerdo del matemático polaco J. M. H. Wronski (1778-1853). Demostrar que la derivada del determinante de $W(x)$ es el determinante de la matriz obtenida derivando cada uno de los elementos de la última fila de $W(x)$.

   [*Indicación:* Utilizar el ejercicio 7.]

4

AUTOVALORES Y AUTOVECTORES

4.1 Transformaciones lineales representadas mediante matrices diagonales

Sea $T: V \to V$ una transformación lineal en un espacio lineal V de dimensión finita. Las propiedades de T que son independientes de cualquier sistema de coordenadas (base), se denominan *propiedades intrínsecas* de T. Si se puede elegir una base de manera que la matriz resultante tenga una forma especialmente sencilla, existe la posibilidad de reconocer algunas de las propiedades directamente a partir de la representación matricial.

Entre los tipos sencillos de matrices están las diagonales. Cabe entonces preguntarse si toda transformación lineal puede representarse mediante una matriz diagonal. En el capítulo 2 se trató el problema de hallar una representación en matriz diagonal de una transformación lineal $T: V \to W$, siendo $\dim V = n$ y $\dim W = m$. En el teorema 2.14 se demostró que siempre existe una base $(e_1, \ldots, e_n)$ para V y otra $(w_1, \ldots, w_m)$ para W tales que la matriz de T en relación a este par de bases es una matriz diagonal. En particular, si $W = V$ la matriz será cuadrada. Ahora se trata de utilizar la *misma base* para V y W. Con esta restricción no siempre es posible encontrar una representación de T en matriz diagonal. Abordemos, entonces, el problema de determinar qué transformaciones tienen una representación en matriz diagonal.

Notación: Si $A = (a_{ij})$ es una matriz diagonal escribimos $A = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$.

Es sencillo establecer una condición necesaria y suficiente para que una transformación lineal tenga una representación en matriz diagonal.

TEOREMA 4.1. *Dada una transformación* $T: V \to V$, *donde* $\dim V = n$. *Si T tiene una representación en matriz diagonal, existe entonces un conjunto de*

elementos independientes $u_1, \ldots, u_n$ *de V y un correspondiente conjunto de escalares* $\lambda_1, \ldots, \lambda_n$ *tales que*

$$(4.1) \qquad T(u_k) = \lambda_k u_k \quad para \quad k = 1, 2, \ldots, n.$$

Recíprocamente, si existe un conjunto independiente $u_1, \ldots, u_n$ *en V y un conjunto correspondiente de escalares* $\lambda_1, \ldots, \lambda_n$ *que satisfacen* (4.1), *entonces la matriz*

$$A = \text{diag}(\lambda_1, \ldots, \lambda_n)$$

es una representación de T respecto a la base $(u_1, \ldots, u_n)$.

Demostración. Supongamos primero que T tiene una representación en matriz diagonal $A = (a_{ik})$ con respecto a la base $(e_1, \ldots, e_n)$. La acción de T sobre los elementos de la base viene dada por la fórmula

$$T(e_k) = \sum_{i=1}^{n} a_{ik} e_i = a_{kk} e_k$$

puesto que $a_{ik} = 0$ para $i \neq k$. Esto demuestra (4.1) con $u_k = e_k$ y $\lambda_k = a_{kk}$.

Supongamos ahora que existen los elementos independientes $u_1, \ldots, u_n$ y los escalares $\lambda_1, \ldots, \lambda_n$ que satisfacen (4.1). Puesto que $u_1, \ldots, u_n$ son independientes, constituyen una base para V. Si definimos $a_{kk} = \lambda_k$ y $a_{ik} = 0$ para $i \neq k$, entonces la matriz $A = (a_{ik})$ es una matriz diagonal que representa T con respecto a la base $(u_1, \ldots, u_n)$.

Así pues, el problema de hallar una representación en matriz diagonal de una transformación lineal se ha reducido a otro, el de hallar los elementos independientes $u_1, \ldots, u_n$ y los escalares $\lambda_1, \ldots, \lambda_n$ que satisfagan (4.1). Tales elementos u_k y escalares λ_k se llaman *autovectores* o *vectores propios* y *autovalores de T*, respectivamente. En la sección siguiente estudiamos los autovectores y autovalores (*) en un aspecto más general.

4.2 Autovectores y autovalores de una transformación lineal

En esta discusión, V representa un espacio lineal y S un subespacio de V. Los espacios S y V no han de ser necesariamente de dimensión finita.

(*) Las palabras *autovector* y *autovalor* son las traducciones de las palabras alemanas *Eigenvektor* y *Eigenwert*. Algunos autores emplean los términos *vector característico*, o *vector propio* como sinónimos de autovector. Los autovalores también se llaman *valores característicos* o *propios*.

DEFINICIÓN. *Sea $T: S \to V$ una transformación de S en V. Un escalar λ se denomina autovalor de T si existe un elemento no nulo x en S tal que*

$$T(x) = \lambda x. \tag{4.2}$$

El elemento x se llama autovector de T perteneciente a λ. El escalar λ se llama autovalor correspondiente a x.

Existe un solo autovalor correspondiente a un autovector dado x. En efecto, si tenemos $T(x) = \lambda x$ y $T(x) = \mu x$ para un cierto $x \neq O$, entonces $\lambda x = \mu x$ con lo que $\lambda = \mu$.

Observación: Si bien la ecuación (4.2) siempre es válida para $x = O$ y cualquier escalar λ, la definición excluye O como autovector. Esto se justifica por tener que existir un solo autovalor λ asociado a un autovector dado x.

Los ejemplos siguientes ponen de manifiesto el significado de esos conceptos.

EJEMPLO 1. *Multiplicación por un escalar fijado.* Sea $T: S \to V$ la transformación lineal definida por la ecuación $T(x) = cx$ para cada x de S, siendo c un escalar fijado. En este ejemplo todo elemento no nulo de S es un autovector perteneciente al escalar c.

EJEMPLO 2. *El autoespacio $E(\lambda)$ que consta de todos los x tales que $T(x) = \lambda x$.* Sea $T: S \to V$ una transformación lineal que tiene un autovalor λ. Sea $E(\lambda)$ el conjunto de *todos* los elementos x de S tales que $T(x) = \lambda x$. Este conjunto contiene el elemento O y todos los autovectores pertenecientes a λ. Es fácil demostrar que $E(\lambda)$ es un subespacio de S, porque si x e y están en $E(\lambda)$ tenemos

$$T(ax + by) = aT(x) + bT(y) = a\lambda x + b\lambda y = \lambda(ax + by)$$

para todos los escalares a y b. Luego $(ax + by) \in E(\lambda)$ así que $E(\lambda)$ es un subespacio. El espacio $E(\lambda)$ se llama *autoespacio* correspondiente a λ. Puede ser de dimensión finita o infinita. Si $E(\lambda)$ es de dimensión finita entonces $\dim E(\lambda) \geq 1$, ya que $E(\lambda)$ contiene por lo menos un elemento x no nulo correspondiente a λ.

EJEMPLO 3. *Existencia de autovalores cero.* Si existe un autovector no puede ser cero, por definición. Sin embargo, el escalar cero *puede* ser un autovalor. En efecto, si 0 es un autovalor para x entonces $T(x) = 0x = O$, por lo que x es del núcleo de T. Recíprocamente, si el núcleo de T contiene elementos no nulos cada uno de ésos es un autovector con autovalor 0. En general, $E(\lambda)$ es el núcleo de $T - \lambda I$.

EJEMPLO 4. *Reflexión en el plano xy.* Sea $S = V = V_3(\mathbf{R})$ y sea T una

reflexión en el plano xy. Esto es, T actúa sobre los vectores bases $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ de la siguiente manera: $T(\boldsymbol{i}) = \boldsymbol{i}$, $T(\boldsymbol{j}) = \boldsymbol{j}$, $T(\boldsymbol{k}) = -\boldsymbol{k}$. Cada vector no nulo en el plano xy es un autovector con un autovalor 1. Los autovectores restantes son de la forma $c\boldsymbol{k}$, donde $c \neq 0$; cada uno de ellos tiene el autovalor -1.

EJEMPLO 5. *Rotación del plano.* Este ejemplo es particularmente interesante porque pone de manifiesto que la existencia de autovectores puede depender del cuerpo fundamental de escalares. El plano puede considerarse como un espacio lineal de dos maneras distintas: 1) como un espacio lineal *real* de dimensión 2, $V = V_2(\mathbf{R})$, con dos elementos base $(1, 0)$ y $(0, 1)$, y con números reales como escalares; o 2) como un espacio lineal *complejo* de dimensión 1, $V = V_1(\mathbf{C})$, con un elemento base 1, y los números complejos como escalares.

Consideremos primero la segunda interpretación. Cada elemento $z \neq 0$ de $V_1(\mathbf{C})$ se puede expresar en forma polar, $z = re^{i\theta}$. Si T imprime a z un giro de ángulo α entonces $T(z) = re^{i(\theta+\alpha)} = e^{i\alpha}z$. Así pues, cada $z \neq 0$ es un autovector con autovalor $\lambda = e^{i\alpha}$. Obsérvese que el autovalor $e^{i\alpha}$ no es real a menos que x sea un múltiplo entero de π.

Consideremos ahora el plano como un espacio lineal *real*, $V_2(\mathbf{R})$. Puesto que los escalares de $V_2(\mathbf{R})$ son números reales, la rotación T admite autovalores reales, únicamente si α es un múltiplo entero de π. Dicho de otro modo, si α no es múltiplo entero de π entonces T no tiene autovalores reales y por tanto tampoco autovectores. De esto resulta claro que la existencia de autovectores puede depender de la elección de los escalares para V.

EJEMPLO 6. *El operador derivada.* Sea V el espacio lineal de todas las funciones reales f que admiten derivadas de cualquier orden en un intervalo abierto dado. Sea D la transformación lineal que aplica cada f en su derivada, $D(f) = f'$. Los autovectores de D son aquellas funciones f no nulas que satisfacen una ecuación de la forma

$$f' = \lambda f$$

para un cierto número real λ. Ésta es una ecuación diferencial lineal de primer orden. Todas sus soluciones vienen dadas por la fórmula

$$f(x) = ce^{\lambda x},$$

donde c es una constante real arbitraria. Por consiguiente, los autovectores de D son todas las funciones exponenciales $f(x) = ce^{\lambda x}$ con $c \neq 0$. El autovalor correspondiente a $f(x) = ce^{\lambda x}$ es λ. En ejemplos como éste, en los que V es un espacio funcional, los autovectores se llaman *autofunciones.*

EJEMPLO 7. *El operador integración.* Sea V el espacio lineal de todas las funciones continuas en un intervalo finito $[a, b]$. Si $f \in V$, designemos por $g = T(f)$ a la función dada por

$$g(x) = \int_a^x f(t)\, dt \qquad \text{si} \quad a \leq x \leq b .$$

Las autofunciones de T (si existen) son aquellas f no nulas que satisfacen una ecuación de la forma

$$\int_a^x f(t)\, dt = \lambda f(x) \tag{4.3}$$

para algún valor real λ. Si existe una autofunción podemos derivar esta ecuación para obtener la relación $f(x) = \lambda f'(x)$, de la cual encontramos $f(x) = ce^{x/\lambda}$, con tal que $\lambda \neq 0$. Dicho de otro modo, las únicas autofunciones posibles son aquellas funciones exponenciales de la forma $f(x) = ce^{x/\lambda}$ con $c \neq 0$ y $\lambda \neq 0$. No obstante, si ponemos $x = a$ en (4.3) obtenemos

$$0 = \lambda f(a) = \lambda c e^{a/\lambda}.$$

Puesto que $e^{a/\lambda}$ nunca es cero, vemos que la ecuación $T(f) = \lambda f$ no puede satisfacerse con una f no nula, así que T no tiene autofunciones ni autovalores.

EJEMPLO 8. *Subespacio engendrado por un autovector.* Sea $T: S \to V$ una transformación lineal que tiene un autovalor λ. Sea x un autovector perteneciente a λ y sea $L(x)$ el subespacio engendrado por x. Esto es, $L(x)$ es el conjunto de todos los productos de x por escalares. Es fácil demostrar que T aplica $L(x)$ en sí mismo. En efecto, si $y = cx$ tenemos

$$T(y) = T(cx) = cT(x) = c(\lambda x) = \lambda(cx) = \lambda y .$$

Si $c \neq 0$ entonces $y \neq O$ por lo que todo elemento no nulo y de $L(x)$ es también un autovector perteneciente a λ.

Se dice que un subespacio U de S es *invariante* en la transformación T, si ésta aplica cada elemento de U en otro elemento de U. Acabamos de demostrar que el subespacio engendrado por un autovector es invariante en T.

4.3 Independencia lineal de autovectores correspondientes a autovalores distintos

Una de las propiedades más importantes de los autovalores es la que se expone en el teorema siguiente. Como antes, S representa un subespacio de un espacio lineal V.

TEOREMA 4.2. *Si $u_1, \ldots, u_k$ son autovectores de una transformación lineal $T: S \to V$, y los autovalores $\lambda_1, \ldots, \lambda_k$ son distintos, entonces los autovectores $u_1, \ldots, u_k$ son independientes.*

Demostración. La demostración es por inducción en k. El resultado es trivial cuando $k = 1$. Supongamos, entonces, que ha sido demostrado para cualquier conjunto de $k-1$ autovectores. Sean $u_1, \ldots, u_k$, k autovectores pertenecientes a autovalores distintos, y supongamos que existen escalares c_i tales que

$$\sum_{i=1}^{k} c_i u_i = O. \tag{4.4}$$

Aplicando T a los dos miembros de (4.4) y utilizando el hecho de que $T(u_i) = \lambda_i u_i$ encontramos

$$\sum_{i=1}^{k} c_i \lambda_i u_i = O. \tag{4.5}$$

Multiplicando (4.4) por λ_k y restando de (4.5) obtenemos la ecuación

$$\sum_{i=1}^{k-1} c_i(\lambda_i - \lambda_k) u_i = O.$$

Pero ya que $u_i, \ldots, u_{k-1}$ son independientes, debe ser $c_i(\lambda_i - \lambda_k) = 0$ para cada $i = 1, 2, \ldots, k-1$. Puesto que los autovalores son distintos tenemos $\lambda_i \neq \lambda_k$ para $i \neq k$ así que $c_i = 0$ para $i = 1, 2, \ldots, k-1$. En virtud de (4.4) vemos que c_k es también 0, por lo que los autovectores $u_1, \ldots, u_k$ son independientes

Obsérvese que el teorema 4.2 no sería cierto si el elemento cero pudiera ser un autovector. Ésta es otra razón para excluir O como autovector.

Advertencia: El recíproco del teorema 4.2 no es cierto. Esto es, si T tiene autovectores independientes $u_1, \ldots, u_k$, entonces los correspondientes autovalores $\lambda_1, \ldots, \lambda_k$ no son necesariamente distintos. Por ejemplo, si T es la transformación idéntica, $T(x) = x$ para todo x, entonces todo $x \neq O$ es un autovector, pero existe sólo un autovalor, $\lambda = 1$.

El teorema 4.2 tiene consecuencias importantes en el caso de dimensión finita.

TEOREMA 4.3. *Si $\dim V = n$, toda transformación lineal $T: V \to V$ tiene por lo menos n autovalores distintos. Si T tiene exactamente n autovalores distintos, entonces los autovectores correspondientes forman una base para V y la*

matriz de T relativa a esa base es una matriz diagonal con los autovalores como elementos diagonales.

Demostración. Si existieran $n + 1$ autovalores distintos, entonces, según el teorema 4.2, V contendría $n + 1$ elementos independientes. Esto no es posible puesto que $\dim V = n$. La segunda afirmación se deduce de los teoremas 4.1 y 4.2.

Observación: El teorema 4.3 nos dice que la existencia de n autovalores distintos es una condición *suficiente* para que T tenga una representación en matriz diagonal. Esta condición no es necesaria. Existen transformaciones lineales con menos de n autovalores distintos que pueden representarse mediante matrices diagonales. La transformación identidad es un ejemplo. Todos sus autovalores son iguales a 1 pero puede representarse por la matriz identidad. El teorema 4.1 nos dice que la existencia de n autovalores independientes es *necesaria y suficiente* para que T admita una representación en matriz diagonal.

4.4 Ejercicios

1. a) Si T tiene un autovalor λ, demostrar que aT tiene el autovalor $a\lambda$.
   b) Si x es un autovector para T_1 y T_2 demostrar que también lo es para $aT_1 + bT_2$. ¿Cómo están relacionados los autovalores?
2. Supongamos que $T: V \to V$ tiene un autovector x perteneciente a un autovalor λ. Demostrar que x es un autovector de T^2 perteneciente a λ^2 y, con mayor generalidad, x es un autovector de T^n perteneciente a λ^n. Utilizar luego el resultado del ejercicio 1 para demostrar que si P es un polinomio, entonces x es un autovector de $P(T)$ perteneciente a $P(\lambda)$.
3. Consideremos el plano como un espacio lineal real, $V = V_2(\mathbf{R})$, y sea T un giro de V de $\pi/2$ radianes. Si bien T no tiene autovectores, demostrar que todo vector no nulo es un autovector de T^2.
4. Si $T: V \to V$ tiene la propiedad de que T^2 tiene un autovalor no negativo λ^2, demostrar que por lo menos uno de los dos valores λ o $-\lambda$ es un autovalor para T.
   [*Indicación:* $T^2 - \lambda^2 I = (T + \lambda I)(T - \lambda I)$.]
5. Sea V el espacio lineal de todas las funciones reales derivables en $(0, 1)$. Si $f \in V$, definamos $g = T(f)$ poniendo $g(t) = tf'(t)$ para todo t en $(0, 1)$. Demostrar que todo λ real es un autovalor de T, y determinar las autofunciones correspondientes a λ.
6. Sea V el espacio lineal de todos los polinomios reales $p(x)$ de grado $\leq n$. Si $p \in V$, definamos $q = T(p)$ poniendo $q(t) = p(t + 1)$ para todo t real. Demostrar que T tiene solamente el autovalor 1. ¿Cuáles son las autofunciones pertenecientes a ese autovalor?
7. Sea V el espacio lineal de todas las funciones continuas en $(-\infty, +\infty)$ y tales que la integral $\int_{-\infty}^{x} f(t)dt$ existe para todo x real. Si $f \in V$ definamos $g = T(f)$ poniendo $g(x) = \int_{-\infty}^{x} f(t)dt$. Demostrar que todo λ positivo es un autovalor para T y determinar las autofunciones correspondientes a λ.
8. Sea V el espacio lineal de todas las funciones continuas en $(-\infty, +\infty)$ y tales que la integral $\int_{-\infty}^{x} t\,f(t)dt$ exista para todo x real. Si $f \in V$ definamos $g = T(f)$ poniendo $g(x) = \int_{-\infty}^{x} tf(t)dt$. Demostrar que todo λ negativo es un autovalor para T y determinar las autofunciones correspondientes a λ.

9. Sea $V = C(0, \pi)$ el espacio lineal real de todas las funciones reales continuas en el intervalo $[0, \pi]$. Sea S el subespacio de todas las funciones f que admiten derivada segunda continua y que satisfacen también las condiciones de contorno $f(0) = f(\pi) = 0$. Sea $T: S \to V$ la transformación lineal que aplica cada f en su derivada segunda, $T(f) = f''$. Demostrar que los autovalores de T son los números de la forma $-n^2$, donde $n = 1, 2, \ldots$, y que las autofunciones correspondientes a $-n^2$ son $f(t) = c_n \operatorname{sen} nt$, siendo $c_n \neq 0$.
10. Sea V el espacio lineal de todas las sucesiones convergentes $\{x_n\}$. Definamos $T: V \to V$ del modo siguiente: Si $x = \{x_n\}$ es una sucesión convergente con límite a, $T(x) = \{y_n\}$, siendo $y_n = a - x_n$ para $n \geq 1$. Demostrar que T tiene sólo dos autovalores, $\lambda = 0$ y $\lambda = -1$, y determinar los autovectores pertenecientes a cada uno de esos λ.
11. Supongamos que una transformación lineal T tiene dos autovectores x e y pertenecientes a autovalores distintos λ y μ. Si $ax + by$ es un autovector de T, demostrar que $a = 0$ o $b = 0$.
12. Sea $T: S \to V$ una transformación lineal tal que todo elemento no nulo de S es un autovector. Demostrar que existe un escalar c tal que $T(x) = cx$. Dicho de otro modo, la única transformación con esta propiedad es el producto de la identidad por un escalar. [*Indicación:* Hacer uso del ejercicio 11.]

4.5 Caso de dimensión finita. Polinomios característicos

Si $\dim V = n$, el problema de determinar los autovalores de una transformación lineal $T: V \to V$ puede resolverse con la ayuda de los determinantes. Deseamos hallar aquellos escalares λ tales que la ecuación $T(x) = \lambda x$ tenga una solución $x \neq O$. La ecuación $T(x) = \lambda x$ puede escribirse en la forma

$$(\lambda I - T)(x) = O,$$

donde I es la transformación identidad. Si ponemos $T_\lambda = \lambda I - T$, entonces λ es un autovalor si y sólo si la ecuación

$$T_\lambda(x) = O \tag{4.6}$$

tiene una solución x no nula, en cuyo caso T_λ no es invertible (en virtud del teorema 2.10). Por consiguiente, según el teorema 2.20, una solución no nula de (4.6) existe si y sólo la matriz de T_λ es singular. Si A es una representación matricial de T, entonces $\lambda I - A$ es una representación matricial para T_λ. Según el teorema 3.13, la matriz $\lambda I - A$ es singular si y sólo si $\det(\lambda I - A) = 0$. Así pues, si λ es un autovalor para T debe satisfacer la ecuación

$$\det(\lambda I - A) = 0. \tag{4.7}$$

Recíprocamente, cualquier λ del *cuerpo fundamental de escalares* que satisfaga (4.7) es un autovalor. Esto sugiere que el determinante $\det(\lambda I - A)$ debería estudiarse como una función de λ.

TEOREMA 4.4. *Si A es una matriz $n \times n$ cualquiera e I es la matriz identidad $n \times n$, entonces la función f definida por la ecuación*

$$f(\lambda) = \det(\lambda I - A)$$

es un polinomio en λ de grado n. Además, el término de mayor grado es λ^n, y el término independiente o constante es $f(0) = \det(-A) = (-1)^n \det A$.

Demostración. La afirmación $f(0) = \det(-A)$ resulta inmediatamente de la definición de f. Vamos a demostrar que f es un polinomio de grado n únicamente para el caso $n \leq 3$. La demostración en el caso general puede hacerse por inducción y la dejamos como ejercicio. (Ver ejercicio 9 de la sección 4.8.)

Para $n = 1$ el determinante es el polinomio lineal $f(\lambda) = \lambda - a_{11}$. Para $n = 2$ tenemos

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{vmatrix} = (\lambda - a_{11})(\lambda - a_{22}) - a_{12}a_{21}$$

$$= \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}),$$

polinomio de segundo grado en λ. Para $n=3$ tenemos

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix}$$

$$= (\lambda - a_{11}) \begin{vmatrix} \lambda - a_{22} & -a_{23} \\ -a_{32} & \lambda - a_{33} \end{vmatrix} + a_{12} \begin{vmatrix} -a_{21} & -a_{23} \\ -a_{31} & \lambda - a_{33} \end{vmatrix} - a_{13} \begin{vmatrix} -a_{21} & \lambda - a_{22} \\ -a_{31} & -a_{32} \end{vmatrix}.$$

Los dos últimos términos son polinomios de primer grado en λ. El primer término es un polinomio de tercer grado, siendo el término de mayor grado λ^3.

DEFINICIÓN. *Si A es una matriz $n \times n$ el determinante*

$$f(\lambda) = \det(\lambda I - A)$$

se denomina polinomio característico de A.

Las raíces del polinomio característico de A son números complejos, alguno de los cuales puede ser real. Si designamos con F el campo real $\mathbf{R}$ o el campo complejo $\mathbf{C}$, tenemos el teorema siguiente.

TEOREMA 4.5. Sea $T: V \to T$ *una transformación lineal, siendo F el campo de escalares de V, y* $\dim V = n$. *Sea A una representación matricial de T. Entonces el conjunto de autovalores de T consta de aquellas raíces del polinomio característico de A que pertenecen a F.*

Demostración. La discusión del teorema 4.4 pone de manifiesto que todo autovalor de T satisface la ecuación $\det(\lambda I - A) = 0$ y que cualquier raíz del polinomio característico de A que pertenece a F es un autovalor de T.

La matriz A depende de la elección de la base para V, pero los autovalores de T se definieron sin hacer referencia a una base. Por consiguiente, el *conjunto* de raíces del polinomio característico de A debe ser independiente de la elección de la base. Más aún, en una sección posterior demostraremos que el mismo polinomio característico es independiente de la elección de base. Volvamos ahora al problema de calcular los autovalores y autovectores en el caso de dimensión finita.

4.6 Cálculo de autovalores y autovectores en el caso de dimensión finita

En el caso de dimensión finita los autovalores y autovectores de una transformación lineal T se llaman también autovalores y autovectores de la representación matricial de T. Así pues, los autovalores de una matriz cuadrada A son las raíces del polinomio característico $f(\lambda) = \det(\lambda I - A)$. Los autovectores correspondientes a un autovalor λ son aquellos vectores no nulos $X = (x_1, \ldots, x_n)$ considerados como matrices columna $n \times 1$ que satisfacen la ecuación matricial

$$AX = \lambda X, \qquad \text{o} \qquad (\lambda I - A)X = O.$$

Éste es un sistema de n ecuaciones lineales para los componentes $x_1, \ldots, x_n$. En cuanto se conoce λ podemos obtener los autovectores resolviendo ese sistema. Seguidamente ofrecemos tres ejemplos que presentan aspecto distinto.

EJEMPLO 1. *Una matriz con todos sus autovalores distintos.* La matriz

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

tiene el polinomio característico

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & -1 & -1 \\ -2 & \lambda - 3 & -4 \\ 1 & 1 & \lambda + 2 \end{bmatrix} = (\lambda - 1)(\lambda + 1)(\lambda - 3),$$

con lo que existen tres autovalores distintos: 1, −1 y 3. Para hallar los autovectores correspondiente a $\lambda = 1$ resolvemos el sistema $AX = X$, o

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Esto nos da

$$\begin{aligned} 2x_1 + x_2 + x_3 &= x_1 \\ 2x_1 + 3x_2 + 4x_3 &= x_2 \\ -x_1 - x_2 - 2x_3 &= x_3, \end{aligned}$$

que podemos escribir en la forma

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 + 4x_3 &= 0 \\ -x_1 - x_2 - 3x_3 &= 0. \end{aligned}$$

Sumando las ecuaciones primera y tercera encontramos $x_3 = 0$, y las tres ecuaciones se reducen entonces a $x_1 + x_2 = 0$. Con lo cual, los autovectores correspondientes a $\lambda = 1$ son $X = t(1, -1, 0)$, siendo t cualquier escalar no nulo.

Mediante cálculos parecidos encontramos los autovectores $X = t(0, 1, -1)$ correspondientes a $\lambda = -1$, y $X = t(2, 3, -1)$ correspondientes a $\lambda = 3$, siendo t un escalar cualquiera no nulo. Puesto que los autovalores son distintos los autovectores correspondientes $(1, -1, 0)$, $(0, 1, -1)$ y $(2, 3, -1)$ son independientes. Los resultados pueden resumirse en forma de tabla de la manera siguiente. En la tercera columna se consigna la dimensión del autoespacio $E(\lambda)$.

Autovalor λ	Autovectores	dim $E(\lambda)$
1	$t(1, -1, 0), \quad t \neq 0$	1
−1	$t(0, 1, -1), \quad t \neq 0$	1
3	$t(2, 3, -1), \quad t \neq 0$	1

EJEMPLO 2. *Una matriz con autovalores repetidos.* La matriz

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}$$

tiene el polinomio característico

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & 1 & -1 \\ 0 & \lambda - 3 & 1 \\ -2 & -1 & \lambda - 3 \end{bmatrix} = (\lambda - 2)(\lambda - 2)(\lambda - 4).$$

Los autovalores son 2, 2 y 4. (Citamos el autovalor 2 dos veces para destacar que es una raíz doble del polinomio característico.) Para encontrar los autovectores correspondientes a $\lambda = 2$ resolveremos el sistema $AX = 2X$, que se reduce a

$$\begin{aligned} -x_2 + x_3 &= 0 \\ x_2 - x_3 &= 0 \\ 2x_1 + x_2 + x_3 &= 0. \end{aligned}$$

Éste tiene la solución $x_2 = x_3 = -x_1$ con lo que los autovectores correspondientes a $\lambda = 2$ son $t(-1, 1, 1)$, donde $t \neq 0$. Análogamente encontramos los autovectores $t(1, -1, 1)$ correspondientes al autovalor $\lambda = 4$. La tabla resumen es la siguiente:

Autovalor λ	Autovectores	dim $E(\lambda)$
2, 2	$t(-1, 1, 1)$, $t \neq 0$	1
4	$t(1, -1, 1)$, $t \neq 0$	1

EJEMPLO 3. *Otra matriz con autovalores repetidos.* La matriz

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

tiene el polinomio característico $(\lambda - 1)(\lambda - 1)(\lambda - 7)$. Cuando $\lambda = 7$ el sistema $AX = 7X$ se convierte

$$\begin{aligned} 5x_1 - x_2 - x_3 &= 0 \\ -2x_1 + 4x_2 - 2x_3 &= 0 \\ -3x_1 - 3x_2 + 3x_3 &= 0. \end{aligned}$$

Éste tiene la solución $x_2 = 2x_1$, $x_3 = 3x_1$, así que los autovectores correspondien-

tes a $\lambda = 7$ son $t(1, 2, 3)$, donde $t \neq 0$. Para el autovalor $\lambda = 1$, el sistema $AX = X$ consta de la ecuación

$$x_1 + x_2 + x_3 = 0$$

repetida tres veces. Para resolver esta ecuación podemos tomar $x_1 = a$, $x_2 = b$, donde a y b son arbitrarios, y tomamos entonces $x_3 = -a - b$. Así, todo autovector correspondiente a $\lambda = 1$ tiene la forma

$$(a, b, -a - b) = a(1, 0, -1) + b(0, 1, -1),$$

siendo $a \neq 0$ o $b \neq 0$. Esto significa que los vectores $(1, 0, -1)$ y $(0, 1, -1)$ forman una base para $E(1)$. Luego $\dim E(\lambda) = 2$ cuando $\lambda = 1$. Los resultados pueden resumirse así:

Autovalor	Autovectores	$\dim E(\lambda)$
7	$t(1, 2, 3)$, $t \neq 0$	1
1, 1	$a(1, 0, -1) + b(0, 1, -1)$, a, b no ambos 0.	2

Obsérvese que en este ejemplo existen tres autovectores independientes pero solamente dos autovalores distintos.

4.7 Traza de una matriz

Sea $f(\lambda)$ el polinomio característico de una matriz A, $n \times n$. Designemos las n raíces de $f(\lambda)$ por $\lambda_1, \ldots, \lambda_n$, repitiendo cada raíz múltiple las veces que indique su orden de multiplicidad. Tenemos entonces la factorización

$$f(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n).$$

También podemos escribir $f(\lambda)$ en potencias decrecientes de λ como sigue,

$$f(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0.$$

Comparándolo con la forma factorizada encontramos que el término independiente c_0 y el coeficiente de λ^{n-1} vienen dados por las fórmulas

$$c_0 = (-1)^n \lambda_1 \cdots \lambda_n \qquad \text{y} \qquad c_{n-1} = -(\lambda_1 + \cdots + \lambda_n).$$

Puesto que también tenemos $c_0 = (-1)^n \det A$, vemos que

$$\lambda_1 \cdots \lambda_n = \det A .$$

Esto es, *el producto de las raíces del polinomio característico de A es igual al determinante de A.*

La *suma* de las raíces de $f(\lambda)$ se llama *traza de A*, y se designa con tr A. Así pues, por definición

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i .$$

El coeficiente de λ^{n-1} viene dado por $c_{n-1} = -\operatorname{tr} A$. También podemos calcular este coeficiente partiendo de $f(\lambda)$ en forma de determinante y encontramos

$$c_{n-1} = -(a_{11} + \cdots + a_{nn}) .$$

(En el ejercicio 12 de la sección 4.8 se pide una demostración de esa fórmula.) Las dos fórmulas de c_{n-1} demuestran que

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii} .$$

Esto es, *la traza de A es también igual a la suma de los elementos diagonales de A.*

Puesto que la suma de los elementos diagonales es fácil de calcular, puede utilizarse como una comprobación numérica en los cálculos de autovalores. En los ejercicios que siguen veremos otras propiedades de la traza.

4.8 Ejercicios

Determinar los autovalores y los autovectores de cada una de las matrices de los ejercicios 1 al 3. Asimismo, para cada autovalor λ calcular la dimensión del autoespacio $E(\lambda)$.

1. a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, c) $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, d) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

2. $\begin{bmatrix} 1 & a \\ b & 1 \end{bmatrix}$, $a > 0, b > 0$.

3. $\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$.

4. Las matrices $P_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $P_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $P_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ se presentan en la

teoría cuántica del spin electrón y se denominan *matrices spin de Pauli*, en honor del físico Wolfgang Pauli (1900-1958). Comprobar que tienen los mismos autovalores 1 y -1. Determinar entonces todas las matrices 2×2 con elementos complejos que tengan los dos autovalores 1 y -1.

5. Determinar todas las matrices 2×2 con elementos reales cuyos autovalores sean a) reales y distintos, b) reales e iguales, c) complejos conjugados.
6. Determinar a, b, c, d, e, f sabiendo que los vectores (1, 1, 1), 1, 0, -1) y (1, -1, 0) son autovectores de la matriz

$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}.$$

7. Calcular los autovalores y autovectores de cada una de las matrices siguientes. También, calcular la dimensión del autoespacio $E(\lambda)$ para cada autovalor λ.

$$\text{a)} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -7 & 1 \end{bmatrix}, \quad \text{b)} \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{bmatrix}, \quad \text{c)} \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

8. Calcular los autovalores de cada una de las cinco matrices

$$\text{a)} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \text{b)} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \text{c)} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$\text{d)} \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \quad \text{e)} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Estas se llaman *matrices de Dirac* en honor de Pablo A. M. Dirac (1902-), físico inglés. Se presentan en la resolución de la ecuación de ondas relativista en Mecánica Cuántica.

9. Si A y B son matrices $n\times n$, siendo B una matriz diagonal, demostrar (por inducción) que el determinante $f(\lambda) = \det(\lambda B - A)$ es un polinomio en λ con $f(0) = (-1)^n \det A$, y con el coeficiente de λ^n igual al producto de los elementos diagonales de B.
10. Demostrar que una matriz cuadrada A y su transpuesta A^t tienen el mismo polinomio característico.
11. Si A y B son matrices $n\times n$ siendo A no singular, demostrar que AB y BA tienen el mismo conjunto de autovalores, incluso si A es singular, pero no es necesario demostrar esto último.

12. Sea A una matriz $n \times n$ con polinomio característico $f(\lambda)$. Demostrar (por inducción) que el coeficiente de λ^{n-1} en $f(\lambda)$ es $-\operatorname{tr} A$.
13. Sean A y B dos matrices $n \times n$ tales que $\det A = \det B$ y $\operatorname{tr} A = \operatorname{tr} B$. Demostrar que A y B tienen el mismo polinomio característico si $n = 2$ pero que esto no es seguro si $n > 2$.
14. Demostrar cada una de las siguientes proposiciones relativas a la traza.

a) $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$.
b) $\operatorname{tr}(cA) = c \operatorname{tr} A$.
c) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.
d) $\operatorname{tr} A^t = \operatorname{tr} A$.

4.9 Matrices que representan la misma transformación lineal. Matrices lineales

En esta sección demostramos que dos representaciones matriciales distintas de una transformación lineal tienen el mismo polinomio característico. Para ello estudiamos con mayor cuidado la relación entre las matrices que representan la misma transformación.

Recordemos cómo se definen las representaciones matriciales. Supongamos que $T: V \to W$ sea una aplicación de un espacio n-dimensional V en un espacio m-dimensional W. Sean $(e_1, \ldots, e_n)$ y $(w_1, \ldots, w_m)$ bases ordenadas para V y W respectivamente. La representación matricial de T relativa a las bases elegidas es la matriz $m \times n$ cuyas columnas constan de los componentes de $T(e_1), \ldots, T(e_n)$ respecto a la base $(w_1, \ldots, w_m)$. Partiendo de bases distintas se obtienen representaciones matriciales distintas.

Consideremos ahora el caso en el que $V = W$, y supongamos que se utiliza la misma base $(e_1, \ldots, e_n)$ para V y para W. Sea $A = (a_{ik})$ la matriz de T relativa a esta base. Esto significa que

$$(4.8) \qquad T(e_k) = \sum_{i=1}^{n} a_{ik} e_i \quad \text{para} \quad k = 1, 2, \ldots, n.$$

Elijamos ahora otra base $(u_1, \ldots, u_n)$ para V y W y sea $B = (b_{kj})$ la matriz de T relativa a esta nueva base. Entonces tenemos

$$(4.9) \qquad T(u_j) = \sum_{k=1}^{n} b_{kj} u_k \quad \text{para} \quad j = 1, 2, \ldots, n.$$

Ya que cada u_j pertenece al espacio engendrado por $e_1, \ldots, e_n$ podemos escribir

(4.10) $$u_j = \sum_{k=1}^{n} c_{kj} e_k \quad \text{para} \quad j = 1, 2, \ldots, n,$$

para un cierto conjunto de escalares c_{kj}. La matriz $n \times n$ $C = (c_{kj})$ determinada por esos escalares es no singular puesto que representa una transformación lineal que aplica una base de V en otra base de V. Aplicando T a los dos miembros de (4.10) también obtenemos las ecuaciones

(4.11) $$T(u_j) = \sum_{k=1}^{n} c_{kj} T(e_k) \quad \text{para} \quad j = 1, 2, \ldots, n.$$

Los sistemas de ecuaciones del (4.8) al (4.11) pueden escribirse con mayor simplicidad en forma matricial introduciendo matrices cuyos elementos sean vectores. Sean

$$E = [e_1, \ldots, e_n] \qquad \text{y} \qquad U = [u_1, \ldots, u_n]$$

matrices fila $1 \times n$ cuyos elementos son los de la base que se considera. Entonces el conjunto de ecuaciones (4.10) puede escribirse mediante una ecuación matricial única,

(4.12) $$U = EC.$$

Análogamente, si introducimos

$$E' = [T(e_1), \ldots, T(e_n)] \qquad \text{y} \qquad U' = [T(u_1), \ldots, T(u_n)],$$

Las ecuaciones (4.8), (4.9) y (4.11) se convierten respectivamente en

(4.13) $$E' = EA, \qquad U' = UB, \qquad U' = E'C.$$

A partir de (4.12) también se obtiene

$$E = UC^{-1}.$$

Para hallar la relación entre A y B expresamos U' de dos maneras en función de U. De (4.13) tenemos

$$U' = UB$$

y

$$U' = E'C = EAC = UC^{-1}AC.$$

Por lo tanto $UB = UC^{-1}AC$. Pero cada elemento en esta ecuación matricial es una combinación lineal de los vectores base $u_1, \ldots, u_n$. Puesto que los u_i son independientes debe ser

$$B = C^{-1}AC.$$

Con ello, hemos demostrado el siguiente teorema.

TEOREMA 4.6. *Si dos matrices $n \times n$ A y B representan la misma transformación lineal T, existe una matriz no singular C tal que*

$$B = C^{-1}AC.$$

Además, si A es la matriz de T relativa a una base $E = [e_1, \ldots, e_n]$ y B la matriz de T relativa a la base $U = [u_1, \ldots, u_n]$, entonces como C podemos tomar la matriz no singular que relaciona las dos bases a través de la ecuación matricial $U = EC$.

También es cierto el recíproco del teorema 4.6.

TEOREMA 4.7. *Sean A y B dos matrices $n \times n$ relacionadas por una ecuación de la forma $B = C^{-1}AC$, en la que C es una matriz $n \times n$ no singular. Entonces A y B representan la misma transformación lineal.*

Demostración. Elijamos una base $E = [e_1, \ldots, e_n]$ para un espacio n-dimensional V. Sean $u_1, \ldots, u_n$ los vectores determinados por las ecuaciones

$$(4.14) \qquad u_j = \sum_{k=1}^{n} c_{kj} e_k \quad \text{para} \quad j = 1, 2, \ldots, n,$$

en donde los escalares c_{kj} son los elementos de C. Puesto que C es no singular representa una transformación lineal invertible, así que $U = [u_1, \ldots, u_n]$ es también una base para V, y tenemos $U = EC$.

Sea T la transformación lineal que tiene A como representación matricial

respecto a la base E, y sea S la transformación lineal que tiene B como representación matricial relativa a la base U. Tenemos entonces

$$T(e_k) = \sum_{i=1}^{n} a_{ik} e_i \quad \text{para} \quad k = 1, 2, \ldots, n \tag{4.15}$$

y

$$S(u_j) = \sum_{k=1}^{n} b_{kj} u_k \quad \text{para} \quad j = 1, 2, \ldots, n. \tag{4.16}$$

Demostraremos que $S = T$ probando que $T(u_j) = S(u_j)$ para cada j.

Las ecuaciones (4.15) y (4.16) pueden escribirse en forma matricial como sigue,

$$[T(e_1), \ldots, T(e_n)] = EA, \qquad [S(u_1), \ldots, S(u_n)] = UB.$$

Aplicando T a (4.14) se obtiene también la relación $T(u_j) = \sum c_{kj} T(e_k)$, o

$$[T(u_1), \ldots, T(u_n)] = EAC.$$

Pero tenemos

$$UB = ECB = EC(C^{-1}AC) = EAC,$$

que demuestra que $T(u_j) = S(u_j)$ para cada j. Por consiguiente, $T(x) = S(x)$ para cada x de V, con lo que $T = S$. Dicho de otro modo, las matrices A y B representan la misma transformación lineal.

DEFINICIÓN. *Dos matrices $n \times n$ A y B se llaman semejantes si existe una matriz no singular C tal que $B = C^{-1}AC$.*

Los teoremas 4.6 y 4.7 pueden combinarse dándonos el siguiente

TEOREMA 4.8. *Dos matrices $n \times n$ son semejantes si y sólo si representan la misma transformación lineal*

Las matrices semejantes tienen muchas propiedades. Por ejemplo, tienen el mismo determinante puesto que

$$\det (C^{-1}AC) = \det (C^{-1})(\det A)(\det C) = \det A.$$

Esta propiedad nos da el teorema siguiente.

TEOREMA 4.9. *Las matrices semejantes tienen el mismo polinomio característico y por tanto los mismos autovalores.*

Demostración. Si A y B son semejantes existe una matriz no singular C tal que $B = C^{-1}AC$. Por consiguiente, tenemos

$$\lambda I - B = \lambda I - C^{-1}AC = \lambda C^{-1}IC - C^{-1}AC = C^{-1}(\lambda I - A)C.$$

Esto prueba que $\lambda I - B$ y $\lambda I - A$ son semejantes, así que $\det(\lambda I - B) = \det(\lambda I - A)$.

Los teoremas 4.8 y 4.9 prueban que todas las representaciones matriciales de una transformación lineal dada T tienen el mismo polinomio característico. Este polinomio también se llama el polinomio característico de T.

El teorema que sigue es una combinación de los teoremas 4.5, 4.2 y 4.6. En el teorema 4.10, F representa bien el cuerpo real **R** o el cuerpo complejo **C**.

TEOREMA 4.10. *Si $T: V \to V$ es una transformación lineal, siendo F el cuerpo de escalares de V y* dim $V = n$ *y suponemos que el polinomio característico de T tiene n raíces distinta $\lambda_1, \ldots, \lambda_n$ en F, entonces:*

a) *Los autovectores correspondientes $u_1, \ldots, u_n$ forman una base para V.*
b) *La matriz de T relativa a la base $U = [u_1, \ldots, u_n]$ es la matriz diagonal Λ que tiene los autovalores como elementos diagonales:*

$$\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n).$$

c) *Si A es la matriz de T relativa a otra base $E = [e_1, \ldots, e_n]$, entonces*

$$\Lambda = C^{-1}AC,$$

donde C es la matriz no singular que relaciona las dos bases mediante la ecuación

$$U = EC.$$

Demostración. Según el teorema 4.5 cada raíz λ_i es un autovalor. Puesto que existen n raíces distintas, el teorema 4.2 nos dice que los autovectores correspondientes $u_1, \ldots, u_n$ son independientes. Luego forman una base para V. Esto demuestra a). Como $T(u_i) = \lambda_i u_i$, la matriz de T relativa a U es la matriz diagonal Λ, lo que demuestra b). Para demostrar c) utilizamos el teorema 4.6.

Observación: La matriz no singular C del teorema 4.10 se llama *matriz diagonalizante.* Si $(e_1, \ldots, e_n)$ es la base de los vectores coordenados unitarios $(I_1, \ldots, I_n)$, entonces la ecuación $U = EC$ del teorema 4.10 demuestra que la columna k de C consta de los componentes de los autovectores u_k relativos a $(I_1, \ldots, I_n)$.

Si los autovalores de A son distintos, entonces A es semejante a una matriz diagonal. Si los autovalores no son distintos, A podría hacerse semejante a una

matriz diagonal. Esto sucederá si y sólo si existen k autovectores independientes correspondientes a cada autovalor de multiplicidad k. Vamos a ver ejemplos en los ejercicios que siguen.

4.10 Ejercicios

1. Demostrar que las matrices $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ tienen los mismos autovalores pero no son semejantes.
2. Hallar en cada caso una matriz no singular C tal que $C^{-1}AC$ es una matriz diagonal o justificar por qué tal matriz C no existe

(a) $A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}$, (b) $A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$, (c) $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$, (d) $A = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$.

3. Se dan en el plano tres bases. Con respecto a esas bases un punto tiene como componentes (x_1, x_2), (y_1, y_2), y (z_1, z_2) respectivamente. Supongamos que $[y_1, y_2] = [x_1, x_2]A$, $[z_1, z_2] = [x_1, x_2]B$, y $[z_1, z_2] = [y_1, y_2]C$ siendo A, B, C matrices 2×2. Expresar C en función de A y B.
4. En cada caso, demostrar que los autovalores de A no son distintos pero que A tiene tres autovectores independientes. Hallar una matriz no singular C tal que $C^{-1}AC$ sea una matriz diagonal.

$$a)\ A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad b)\ A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -1 & -1 & 1 \end{bmatrix}.$$

5. Demostrar que ninguna de las matrices siguientes es semejante a una matriz diagonal, pero que cada una es semejante a una matriz triangular de la forma $\begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}$ en la que λ es un autovalor.

$$a)\ \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}, \qquad b)\ \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}.$$

6. Determinar los autovalores y los autovectores de la matriz $\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & -3 & 3 \end{bmatrix}$ y con ello demostrar que no es semejante a una matriz diagonal.

7. a) Demostrar que una matriz cuadrada A es no singular si y sólo si 0 no es autovalor de A.

b) Si A es no singular, demostrar que los autovalores de A^{-1} son los recíprocos de los autovalores de A.

8. Dada una matriz A $n \times n$ con elementos reales tal que $A^2 = -I$. Demostrar las proposiciones siguientes referentes a A.
   a) A es no singular.
   b) n es par.
   c) A no tiene autovalores reales.
   d) $\det A = 1$.

5

AUTOVALORES DE OPERADORES EN ESPACIOS EUCLÍDEOS

5.1 Autovalores y productos interiores o escalares

Este capítulo se ocupa de algunas propiedades de los autovalores y autovectores de las transformaciones lineales que operan en espacios euclídeos, esto es, en espacios lineales que tienen producto interior o escalar. Recordemos las propiedades fundamentales de los productos interiores.

En un espacio euclídeo *real* un producto interior (x,y) de dos elementos x e y es un número real que satisface las propiedades siguientes:

(1) $(x, y) = (y, x)$ (simetría)
(2) $(x + z, y) = (x, y) + (z, y)$ (linealidad)
(3) $(cx, y) = c(x, y)$ (homogeneidad)
(4) $(x, x) > 0$ if $x \neq O$ (positividad).

En un espacio euclídeo *complejo* el producto interior es un número complejo que satisface las mismas propiedades, con la excepción de que la simetría se reemplaza por la *simetría de Hermite* o *hermitiana,*

$$(1') \qquad (x, y) = \overline{(y, x)},$$

en donde la barra representa el complejo conjugado. En 3) el escalar c es complejo. De (1′) y (3) obtenemos

$$(3') \qquad (x, cy) = \bar{c}(x, y),$$

que nos dice que los escalares son conjugados cuando se sacan del segundo factor. Haciendo en (1′) $x = y$ vemos que (x, x) es real, con lo que la propiedad (4) tiene sentido si el espacio es complejo.

Cuando hablamos de espacio euclídeo, sin más, se sobreentiende que el espacio puede ser real o complejo. Si bien la mayor parte de nuestras aplicaciones se harán en espacios de dimensión finita, no impondremos esa restricción a priori.

El primer teorema demuestra que los autovalores (si existen) pueden expresarse en función del producto interior.

TEOREMA 5.1. *Si E es un espacio euclídeo, V un subespacio de E y consideramos una transformación lineal $T: V \to E$ que tenga un autovalor λ con un correspondiente autovector x, entonces*

$$\lambda = \frac{(T(x), x)}{(x, x)}. \tag{5.1}$$

Demostración. Puesto que $T(x) = \lambda x$ *tenemos*

$$(T(x), x) = (\lambda x, x) = \lambda(x, x).$$

Como $x \neq O$ podemos dividir por (x, x) obteniendo (5.1).

Varias propiedades de los autovalores se deducen fácilmente a partir de la ecuación (5.1). Por ejemplo, de la simetría hermitiana del producto interior tenemos la fórmula análoga

$$\bar{\lambda} = \frac{(x, T(x))}{(x, x)} \tag{5.2}$$

para el complejo conjugado de λ. De (5.1) y (5.2) deducimos que λ es real ($\lambda = \bar{\lambda}$) si y sólo si $(T(x), x)$ es imaginario puro, esto es, si y sólo si

$$(T(x), x) = (x, T(x)) \qquad \text{para el autovector } x.$$

(Esta condición se cumple en un espacio euclídeo real). También, λ es imaginario puro ($\lambda = -\bar{\lambda}$) si y solamente si $(T(x), x)$ es imaginario puro, lo que es lo mismo, si y solamente si

$$(T(x), x) = -(x, T(x)) \qquad \text{para el autovector } x.$$

5.2 Transformaciones hermitianas y hemi-hermitianas

En esta sección introducimos dos tipos importantes de operadores lineales que actúan en espacios euclídeos. Esos operadores tienen dos categorías de denominaciones, que dependen de si el espacio euclídeo fundamental tiene producto interior real o complejo. En el caso real las transformaciones se llaman *simétricas* o *hemi-simétricas*. En el caso complejo se llaman *hermitianas* o *hemi-hermitianas*. Esas transformaciones se presentan en multitud de aplicaciones distintas. Por

ejemplo, los operadores hermitianos en espacios de dimensión infinita desempeñan un importante papel en Mecánica Cuántica. Discutiremos en primer lugar el caso complejo ya que no presenta nuevas dificultades.

DEFINICIÓN. *Sean E un espacio euclídeo y V un subespacio de E. Una transformación lineal $T: V \to E$ se llama hermitiana en V si*

$$(T(x), y) = (x, T(y)) \qquad \text{para todo } x \text{ y todo } y \text{ de } V.$$

El operador T se llama hemi-hermitiano en V si

$$(T(x), y) = -(x, T(y)) \qquad \text{para todo } x \text{ y todo } y \text{ de } V.$$

Dicho de otro modo, un operador hermitiano T puede pasarse de un factor de un producto interior al otro sin modificar el valor del producto. Tal paso en el caso de un operador hemi-hermitiano cambia el signo del producto.

Observación: Como ya se ha mencionado, si E es un espacio euclídeo *real*, las transformaciones hermitianas también se llaman *simétricas;* las transformaciones hemi-hermitianas se llaman *hemi-simétricas.*

EJEMPLO 1. *Simetría y hemi-simetría en el espacio $C(a, b)$.* Sea $C(a, b)$ el espacio de todas las funciones reales continuas en un intervalo cerrado $[a, b]$, con el producto interior real

$$(f, g) = \int_a^b f(t)g(t)\, dt.$$

Sea V un subespacio de $C(a, b)$. *Si* $T: V \to C(a, b)$ es una transformación lineal, entonces $(f, T(g)) = \int_a^b f(t)Tg(t)dt$, donde hemos escrito $Tg(t)$ en lugar de $T(g)(t)$ Por consiguiente, las condiciones para la simetría y la hemi-simetría se convierten en

$$(5.3) \qquad \int_a^b \{f(t)Tg(t) - g(t)Tf(t)\}\, dt = 0 \qquad \text{si } T \text{ es simétrica,}$$

y

$$(5.4) \qquad \int_a^b \{f(t)Tg(t) + g(t)Tf(t)\}\, dt = 0 \qquad \text{si } T \text{ es hemi-simétrica.}$$

EJEMPLO 2. *Multiplicación por una función fija.* En el espacio $C(a, b)$ del ejemplo 1, elijamos una función fija p y definamos $T(f) = pf$, producto de p y f. Para esa T, la ecuación (5.3) se satisface cualesquiera que sean f y g en $C(a, b)$

puesto que el integrando es cero. Por consiguiente, la multiplicación por una función fija es un operador simétrico.

EJEMPLO 3. *El operador derivada.* En el espacio $C(a, b)$ del ejemplo 1, sea V el subespacio de todas las funciones f que tienen derivada continua en el intervalo abierto (a, b) y que satisfacen también la *condición de contorno* $f(a) = f(b)$. Sea $D: V \to C(a, b)$ el operador derivada dado por $D(f) = f'$. Resulta sencillo demostrar que D es hemi-simétrico. En este caso el integrando de (5.4) es la derivada del producto fg, por lo cual la integral es igual a

$$\int_a^b (fg)'(t)\, dt = f(b)g(b) - f(a)g(a).$$

Puesto que f y g satisfacen ambas la condición de contorno, tenemos $f(b)g(b) - f(a)g(a) = 0$. Así pues, la condición de contorno implica la hemi-simetría de D. Las funciones constantes son las únicas autofunciones en el subespacio V. Pertenecen al autovalor 0.

EJEMPLO 4. *Operadores de Sturm-Liouville.* Este ejemplo es importante en la teoría de las ecuaciones diferenciales lineales de segundo orden. Utilicemos una vez más el espacio $C(a, b)$ del ejemplo 1 y sea V el subespacio de todas las funciones f que tienen derivada segunda continua en $[a, b]$ y que satisfacen las condiciones de contorno

$$p(a)f(a) = 0, \qquad p(b)f(b) = 0, \tag{5.5}$$

siendo p una función fija de $C(a, b)$ con derivada continua en $[a, b]$. Sea q otra función fija en $C(a, b)$ y sea $T: V \to C(a, b)$ el operador definido por la ecuación

$$T(f) = (pf')' + qf.$$

Éste se llama *operador de Sturm-Liouville.* Para averiguar su simetría observemos que $fT(g) - gT(f) = f(pg')' - g(pf')'$. Utilizando este resultado en (5.3) e integrando por partes, encontramos

$$\int_a^b \{fT(g) - gT(f)\}\, dt = fpg' \Big|_a^b - \int_a^b pg'f'\, dt - gpf' \Big|_a^b + \int_a^b pf'g'\, dt = 0,$$

ya que tanto f como g satisfacen las condiciones de contorno (5.5). Por tanto T es simétrico en V. Las autofunciones de T son aquellas funciones no nulas

f que satisfacen, para un cierto valor real λ, una ecuación diferencial de la forma

$$(pf')' + qf = \lambda f$$

en $[a, b]$, y las condiciones de contorno (5.5).

5.3 Autovalores y autovectores de los operadores hermitianos y hemi-hermitianos

En cuanto a los autovalores tenemos el siguiente teorema:

TEOREMA 5.2. *Si T tiene un autovalor λ, entonces:*
a) *Si T es hermitiano, λ es real: $\lambda = \bar{\lambda}$.*
b) *Si T es hemi-hermitiano, λ es imaginario puro: $\lambda = -\bar{\lambda}$.*

Demostración. Sea x un autovector correspondiente a λ. Tenemos entonces

$$\lambda = \frac{(T(x), x)}{(x, x)} \qquad \text{y} \qquad \bar{\lambda} = \frac{(x, T(x))}{(x, x)}.$$

Si T es hermitiano tenemos $(T(x), x) = (x, T(x))$ así que $\lambda = \bar{\lambda}$. Si T es hemi-hermitiano tenemos $(T(x), x) = -(x, T(x))$ con lo cual $\lambda = -\bar{\lambda}$.

Observación: Si T es *simétrico,* el teorema 5.2 nada nuevo nos dice acerca de los autovalores de T puesto que todos los autovalores deben ser reales si su producto interior es real. Si T es *hemi-simétrico,* los autovalores de T deben ser a un tiempo reales e imaginarios puros, luego deben ser cero (si existen).

5.4 Ortogonalidad de los autovectores correspondientes a autovalores distintos

Los autovalores distintos de una transformación lineal cualquiera corresponden a autovectores independientes (según el teorema 4.2). Para las transformaciones hermitianas y hemi-hermitianas podemos decir aún más.

TEOREMA 5.3. *Si T es una transformación hermitiana o hemi-hermitiana y si λ y μ son autovalores distintos de T con los autovectores correspondientes x e y, entonces x e y son ortogonales; esto es $(x, y) = 0$.*

Demostración. Escribamos $T(x) = \lambda x$, $T(y) = \mu y$ y comparemos los dos productos interiores $(T(x), y)$ y $(x, T(y))$. Tenemos

$$(T(x), y) = (\lambda x, y) = \lambda(x, y) \qquad \text{y} \qquad (x, T(y)) = (x, \mu y) = \bar{\mu}(x, y).$$

Si T es hermitiana esto nos da $\lambda(x, y) = \bar{\mu}(x, y) = \mu(x, y)$ ya que $\mu = \bar{\mu}$. Por consiguiente $(x, y) = 0$ puesto que $\lambda \neq \mu$. Si T es hemi-hermitiana obtenemos $\lambda(x, y) = -\bar{\mu}(x, y) = \mu(x, y)$ que implica también $(x, y) = 0$.

EJEMPLO. Apliquemos el teorema 5.3 a aquellas funciones no nulas que satisfacen una ecuación diferencial de la forma

$$(pf')' + qf = \lambda f \tag{5.6}$$

en un intervalo $[a, b]$, y que satisfacen también las condiciones de contorno $p(a)f(a) = p(b)f(b) = 0$. La conclusión es que dos soluciones cualesquiera f y g correspondientes a dos valores distintos de λ son ortogonales. Por ejemplo, consideremos la ecuación diferencial del movimiento armónico simple

$$f'' + k^2 f = 0$$

en el intervalo $[0, \pi]$, siendo $k \neq 0$. Ésta tiene la forma (5.6) con $p = 1$, $q = 0$, y $\lambda = -k^2$. Todas las soluciones vienen dadas por $f(t) = c_1 \cos kt + c_2 \operatorname{sen} kt$. La condición de contorno $f(0) = 0$ implica $c_1 = 0$. La segunda condición de contorno, $f(\pi) = 0$, implica $c_2 \operatorname{sen} k\pi = 0$. Puesto que $c_2 \neq 0$ para una solución no nula, debe ser $\operatorname{sen} k\pi = 0$, lo que indica que k es entero. Dicho de otro modo, las soluciones no nulas que satisfacen las condiciones de contorno existen si y sólo si k es entero. Esas soluciones son $f(t) = \operatorname{sen} nt$, $n = \pm 1, \pm 2, \ldots$. La condición de ortogonalidad que el teorema 5.3 implica, se transforma ahora en la relación conocida

$$\int_0^\pi \operatorname{sen} nt \operatorname{sen} mt \, dt = 0$$

si m^2 y n^2 son enteros distintos.

5.5 Ejercicios

1. Sean E un espacio euclídeo y V un subespacio, y sea $T: V \to E$ una transformación lineal dada. Sean λ un escalar y x un elemento no nulo de V. Demostrar que λ es un autovalor de T, con x como autovector, si y sólo si

$$(T(x), y) = \lambda(x, y) \qquad \text{para todo } y \text{ de } E.$$

2. Sea $T(x) = cx$ para todo x un espacio lineal V, siendo c un escalar fijado. Demostrar que T es simétrico si V es un espacio euclídeo real.
3. Supongamos que $T: V \to V$ es una transformación hermitiana.
   a) Demostrar que T^n es hermitiana para todo entero positivo n, y que T^{-1} es hermitiana si T es invertible.
   b) ¿Qué puede decirse de T^n y de T^{-1} si T es hemi-hermitiana?

4. Sean T_1: $V \to E$ y T_2: $V \to E$ dos transformaciones hermitianas.
   a) Demostrar que $aT_1 + bT_2$ es hermitiana para todo par de escalares reales a y b.
   b) Demostrar que el producto (la composición) T_1T_2 es hermitiana si T_1 y T_2 son permutables, esto es, si $T_1T_2 = T_2T_1$.
5. Sea $V = V_3(\mathbf{R})$ con el producto escalar ordinario como producto interior. Sea T una simetría en el plano xy; esto es, $T(\boldsymbol{i}) = \boldsymbol{i}$, $T(\boldsymbol{j}) = \boldsymbol{j}$ y $T(\boldsymbol{k}) = -\boldsymbol{k}$. Demostrar que T es simétrica.
6. Sea $C(0, 1)$ el espacio lineal real de todas las funciones continuas en $[0, 1]$ con producto interior $(f, g) = \int_0^1 f(t)g(t)dt$. Sea V el subespacio de todas las f tales que $\int_0^1 f(t)dt = 0$. Sea T: $V \to C(0, 1)$ el operador integración definido por $Tf(x) = \int_0^x f(t)dt$. Demostrar que T es hemi-simétrica.
7. Sea V el espacio euclídeo real de todos los polinomios con el producto interior $(f, g) = \int_{-1}^1 f(t)g(t)dt$. Determinar cuáles de las siguientes transformaciones $T: V \to V$ son simétricas o hemi-simétricas:

   a) $Tf(x) = f(-x)$.
   b) $Tf(x) = f(x)f(-x)$.
   c) $Tf(x) = f(x) + f(-x)$.
   d) $Tf(x) = f(x) - f(-x)$.

8. En el ejemplo 4 de la sección 5.2, modificar el producto interior así:

$$(f, g) = \int_a^b f(t)g(t)w(t)\,dt,$$

siendo w una función positiva fijada perteneciente a $C(a, b)$. Modificar el operador T de Sturm-Liouville escribiendo

$$T(f) = \frac{(pf')' + qf}{w}.$$

Demostrar que el operador modificado es simétrico en el subespacio V.

9. Sea V un subespacio de un espacio euclídeo complejo E. Sea T: $V \to E$ una transformación lineal y definamos una función de valores escalares Q en V, como sigue:

$$Q(x) = (T(x), x) \qquad \text{para todo } x \text{ de } V.$$

   a) Si T es hermitiana en V, demostrar que $Q(x)$ es real para todo x.
   b) Si T es hemi-hermitiana, demostrar que $Q(x)$ es imaginaria pura para todo x.
   c) Demostrar que $Q(tx) = t\bar{t}Q(x)$ para todo escalar t.
   d) Demostrar que $Q(x + y) = Q(x) + Q(y) + (T(x), y) + (T(y), x)$, y hallar la fórmula análoga para $Q(x + ty)$.
   e) Si $Q(x) = 0$ para todo x demostrar que $T(x) = O$ para todo x.
   f) Si $Q(x)$ es real para todo x demostrar que T es hermitiana. [*Indicación:* Utilizar la propiedad de que $Q(x + ty)$ es igual a su conjugada para todo escalar t.]
10. Este ejercicio pone de manifiesto que los polinomios de Legendre (introducidos en la sección 1.14) son autofunciones de un operador de Sturm-Liouville. Los polinomios de Legendre se definen mediante la ecuación

$$P_n(t) = \frac{1}{2^n n!} f_n^{(n)}(t), \qquad \text{donde} \quad f_n(t) = (t^2 - 1)^n.$$

   a) Comprobar que $(t^2 - 1)f'_n(t) = 2ntf_n(t)$.

b) Derivar la ecuación del apartado a) $n+1$ veces, mediante la fórmula de Leibniz (ver p. del Volumen I) con lo que se obtiene

$$(t^2-1)f_n^{(n+2)}(t)+2t(n+1)f_n^{(n+1)}(t)+n(n+1)f_n^{(n)}(t)$$
$$=2ntf_n^{(n+1)}(t)+2n(n+1)f_n^{(n)}(t).$$

c) Demostrar que la ecuación del apartado b) puede ponerse en la forma

$$[(t^2-1)P_n'(t)]'=n(n+1)P_n(t).$$

Esto demuestra que $P_n(t)$ es una autofunción del operador de Sturm-Liouville dado en el intervalo $[-1,1]$ por medio de $T(f)=(pf')'$, siendo $p(t)=t^2-1$. La autofunción $P_n(t)$ pertenece al autovalor $\lambda=n(n+1)$. En este ejemplo las condiciones de contorno para la simetría se satisfacen automáticamente puesto que $p(1)=p(-1)=0$.

5.6 Existencia de un conjunto ortonormal de autovectores para operadores hermitianos y hemi-hermitianos que actúan en espacios de dimensión finita

Los dos teoremas 5.2 y 5.3 se fundamentan en la hipótesis de que T tiene un autovalor. Como sabemos, los autovalores no existen necesariamente. No obstante si T actúa en un espacio *complejo* de dimensión *finita,* entonces existen siempre los autovalores puesto que son las raíces del polinomio característico. Si T es hermitiano, todos los autovalores son reales. Si T es hemi-hermitiano, todos los autovalores son imaginarios puros.

También sabemos que dos autovalores distintos pertenecen a autovectores ortogonales, si T es hermitiano o hemi-hermitiano. Usando esta propiedad se puede probar que T tiene un conjunto ortonormal de autovectores que engendran todo el espacio. (Recordemos que un conjunto ortogonal se llama ortonormal si cada uno de sus elementos tiene norma 1.)

TEOREMA 5.4. *Supongamos que* $\dim V=n$ *y sea* $T\colon V\to V$ *hermitiano o hemi-hermitiano. Existen entonces* n *autovectores* $u_1,\ldots,u_n$ *de* T *que forman una base ortonormal para* V. *Por tanto, la matriz de* T *relativa a esta base es la matriz diagonal* $\Lambda=\operatorname{diag}(\lambda_1,\ldots,\lambda_n)$, *en donde* λ_k *es el autovalor perteneciente a* u_k.

Demostración. Utilicemos el método de inducción respecto a la dimensión n. Si $n=1$, entonces T tiene exactamente un autovalor. Cualquier autovector u_1 de norma 1 es una base ortonormal para V.

Supongamos ahora el teorema cierto para todo espacio euclídeo de dimensión $n-1$. Para demostrar que es cierto para V elijamos un autovalor λ_1 para

T y un autovector correspondiente u_1 de norma 1. Entonces $T(u_1) = \lambda_1 u_1$ y $\|u_1\| = 1$. Sea S el subespacio engendrado por u_1 . Aplicaremos la hipótesis de inducción al subespacio $S^\perp$ que consta de todos los elementos de V que son ortogonales a u_1 ,

$$S^\perp = \{x \mid x \in V, (x, u_1) = 0\}.$$

Para ello necesitamos saber que $\dim S^\perp = n-1$ y que T aplica $S^\perp$ en sí mismo.

Según el teorema 1.7 a) sabemos que u_1 es parte de una base para V, sea ésta la base $(u_1, v_2, \ldots, v_n)$. Podemos suponer, sin pérdida de generalidad, que ésa es una base ortonormal. (Si no lo fuera, aplicaríamos el método de Gram-Schmidt para convertirla en tal, manteniendo u_1 como primer elemento de la base.) Tomemos seguidamente un x cualquiera en $S^\perp$ y escribamos

$$x = x_1u_1 + x_2v_2 + \cdots + x_nv_n.$$

Entonces $x_1 = (x, u_1) = 0$ ya que la base es ortonormal, así que x pertenece al espacio engendrado por $v_2, \ldots, v_n$. Luego $\dim S^\perp = n-1$.

Demostremos a continuación que T aplica $S^\perp$ en sí mismo. Supongamos que T es hermitiana. Si $x \in S^\perp$ tenemos

$$(T(x), u_1) = (x, T(u_1)) = (x, \lambda_1 u_1) = \lambda_1(x, u_1) = 0,$$

por lo que $T(x) \in S^\perp$. Puesto que T es hermitiana en $S^\perp$ podemos aplicar la hipótesis de inducción encontrando que T tiene $n-1$ autovectores $u_2, \ldots, u_n$ que constituyen una base ortonormal para $S^\perp$. Por lo tanto, el conjunto ortogonal $u_1, \ldots, u_n$ es una base ortonormal para V. Esto demuestra el teorema si T es hermitiana. Argumento similar aplicaremos si T es hemi-hermitiana.

5.7 Representación matricial para operadores hermitianos y hemi-hermitianos

Suponemos en esta sección que V es un espacio euclídeo de dimensión finita. Una transformación hermitiana o hemi-hermitiana puede caracterizarse por su acción sobre los elementos de una base cualquiera.

TEOREMA 5.5. *Sean $(e_1, \ldots, e_n)$ una base para V y $T\colon V \to V$ una transformación lineal. Tenemos entonces:*

a) *T es hermitiana si y sólo si $(T(e_j), e_i) = (e_j, T(e_i))$ para todo par i, j.*

b) *T es hemi-hermitiana si y sólo si $(T(e_j), e_i) = -(e_j, T(e_i))$ para todo par i, j.*

Demostración. Tomemos dos elementos cualesquiera x e y de V y expre

semos cada uno en función de los elementos de la base, sean $x = \sum x_j e_j$ e $y = \sum y_i e_i$. Tenemos entonces

$$(T(x), y) = \left(\sum_{j=1}^{n} x_j T(e_j), y\right) = \sum_{j=1}^{n} x_j \left(T(e_j), \sum_{i=1}^{n} y_i e_i\right) = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j \bar{y}_i (T(e_j), e_i).$$

Del mismo modo encontramos

$$(x, T(y)) = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j \bar{y}_i (e_j, T(e_i)).$$

Las proposiciones a) y b) se deducen al momento a partir de esas ecuaciones.

Vamos a expresar estos conceptos por medio de la representación matricial de T.

TEOREMA 5.6. *Sea $(e_1, \ldots, e_n)$ una base ortonormal para V, y sea $A = (a_{ij})$ la representación matricial de una transformación lineal $T: V \to V$ respecto de esa base. Tenemos entonces:*

a) *T es hermitiana si y sólo si $a_{ij} = \bar{a}_{ji}$ para todo par i, j.*

b) *T es hemi-hermitiana si y sólo si $a_{ij} = -\bar{a}_{ji}$ para todo par i, j.*

Demostración. Puesto que A es la matriz de T tenemos $T(e_j) = \sum_{k=1}^{n} a_{kj} e_k$. Tomando el producto interior de $T(e_j)$ por e_i y teniendo en cuenta la linealidad del producto interior obtenemos

$$(T(e_j), e_i) = \left(\sum_{k=1}^{n} a_{kj} e_k, e_i\right) = \sum_{k=1}^{n} a_{kj} (e_k, e_i).$$

Pero $(e_k, e_i) = 0$ salvo si $k = i$, así que la última suma se reduce a $a_{ij}(e_i, e_i) = a_{ij}$ ya que $(e_i, e_i) = 1$. Luego resulta

$$a_{ij} = (T(e_j), e_i) \quad \text{para todo par } i, j.$$

Intercambiando i y j, tomando los conjugados, y teniendo en cuenta la simetría hermitiana del producto interior, encontramos

$$\bar{a}_{ji} = (e_j, T(e_i)) \quad \text{para todo par } i, j.$$

Para completar la demostración basta aplicar ahora el teorema 5.5.

5.8 Matrices hermitianas y hemi-hermitianas. Matriz adjunta de una matriz

El teorema 5,6 nos sugiere la siguiente definición.

DEFINICIÓN. *Una matriz cuadrada $A = (a_{ij})$ se denomina hermitiana si $a_{ij} = \bar{a}_{ji}$ para todo par i, j. Se dice que A es hemi-hermitiana si $a_{ij} = -\bar{a}_{ji}$ para todo par i, j.*

El teorema 5.6 establece que una transformación T en un espacio V de dimensión finita es hermitiana o hemi-hermitiana según que su matriz relativa a una base ortonormal sea hermitiana o hemi-hermitiana.

Esas matrices pueden introducirse de otra manera. Llamemos $\bar{A}$ la matriz obtenida reemplazando cada elemento de A por su complejo conjugado. La matriz $\bar{A}$ se llama la *conjugada* de A. La matriz A es hermitiana si y sólo si es igual a la transpuesta de su conjugada $A = \bar{A}^t$. Es hemi-hermitiana si $A = -\bar{A}^t$.

La transpuesta de la conjugada recibe un nombre especial.

DEFINICIÓN DE MATRIZ ADJUNTA DE UNA MATRIZ DADA. *Dada una matriz cualquiera A, la transpuesta de la conjugada, $\bar{A}^t$, se llama también adjunta de A y se representa por A^*.*

Así pues, una matriz cuadrada A es hermitiana si $A = A^*$, y hemi-hermitiana si $A = -A^*$. Una matriz hermitiana se llama también *auto-adjunta.*

Observación: Gran parte de la antigua literatura relativa a matrices utiliza la denominación de *adjunta* para la transpuesta de la matriz cofactor, que es un ente completamente distinto. La definición dada aquí está de acuerdo con la nomenclatura actual de la teoría de operadores lineales.

5.9 Diagonalización de una matriz hermitiana o hemi-hermitiana

TEOREMA 5.7. *Toda matriz A, $n \times n$, hermitiana o hemi-hermitiana es semejante a la matriz diagonal $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ de sus autovalores. Además, tenemos*

$$\Lambda = C^{-1}AC,$$

en donde C es una matriz no singular cuya inversa es su adjunta, $C^{-1} = C^*$.

Demostración. Sea V el espacio de las *n-plas* de números complejos, y sea $(e_1, \ldots, e_n)$ la base ortonormal de vectores coordenados unitarios. Si $x = \sum x_i e_i$ e $y = \sum y_i e_i$, consideremos el producto interior dado por $(x, y) = \sum x_i \bar{y}_i$. Para la matriz dada A, sea T la transformación representada por A relativa a la base elegida. Entonces el teorema 5.4 nos dice que V tiene una base ortonormal de autovectores $(u_1, \ldots, u_n)$, respecto a la cual T tiene una representación en matriz diagonal $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, siendo λ_k el autovalor per-

teneciente a u_k. Puesto que tanto A como Λ representan T, serán semejantes, así que tenemos $\Lambda = C^{-1}AC$, donde $C = (c_{ij})$ es la matriz no singular que relaciona las dos bases:

$$[u_1, \ldots, u_n] = [e_1, \ldots, e_n]C.$$

Esta ecuación muestra que la columna j de C consta de los componentes de u_j respecto a $(e_1, \ldots, e_n)$. Por consiguiente, c_{ij} es el componente i-ésimo de u_j. El producto interior de u_j y u_i viene dado por

$$(u_j, u_i) = \sum_{k=1}^{n} c_{kj}\bar{c}_{ki}.$$

Puesto que $\{u_1, \ldots, u_n\}$ es un conjunto ortonormal, eso demuestra que $CC^* = I$, con lo que $C^{-1} = C^*$.

Observación: La demostración del teorema 5.7 nos dice también la manera de determinar la matriz diagonalizante C. Encontramos un conjunto ortonormal de autovectores $u_1, \ldots, u_n$ y luego utilizamos los componentes de u_j (relativos a la base de vectores coordenados unitarios) como elementos de la columna j de C.

EJEMPLO 1. La matriz real hermitiana $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$ tiene los autovalores $\lambda_1 = 1$ y $\lambda_2 = 6$. Los autovectores pertenecientes a 1 son $t(2, -1)$, $t \neq 0$. Los pertenecientes a 6 son $t(1, 2)$, $t \neq 0$. Los dos autovectores $u_1 = t(2, -1)$ y $u = t(1, 2)$ con $t = 1/\sqrt{5}$ constituyen un conjunto ortonormal. Por consiguiente, la matriz

$$C = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$$

es una matriz diagonalizante para A. En este caso $C^* = C^t$ puesto que C es real.

Fácilmente se comprueba que $C^tAC = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$.

EJEMPLO 2. Si A es ya una matriz diagonal, la matriz diagonalizante C del teorema 5.7 o bien deja invariable A o tan sólo reordena los elementos diagonales.

5.10 Matrices unitarias. Matrices ortogonales

DEFINICIÓN. *Una matriz cuadrada A se llama unitaria si $AA^* = I$. Se llama ortogonal si $AA^t = I$.*

Observación: Toda matriz unitaria real es ortogonal puesto que $A^* = A^t$.

El teorema 5.7 nos dice que una matriz hermitiana o hemi-hermitiana siempre puede ser diagonalizada por medio de una matriz unitaria. Una matriz real hermitiana tiene autovalores reales y los autovectores correspondientes pueden tomarse reales. Por consiguiente, una matriz real hermitiana puede ser diagonalizada por una matriz real ortogonal. Esto *no* es cierto para las matrices hemi-hermitianas reales. (Ver el ejercicio 11 de la sección 5.11.)

Asimismo tenemos los conceptos relacionados siguientes,

DEFINICIÓN. *Una matriz cuadrada* A *con elementos reales o complejos se llama simétrica si* $A = A^t$; *se llama hemi-simétrica si* $A = -A^t$.

EJEMPLO 3. Si A es real, su adjunta es igual a su transpuesta, $A^* = A^t$. Así pues, toda matriz hermitiana *real* es simétrica, pero una matriz simétrica no es necesariamente hermitiana.

EJEMPLO 4. Si $A = \begin{bmatrix} 1+i & 2 \\ 3-i & 4i \end{bmatrix}$, entonces $\bar{A} = \begin{bmatrix} 1-i & 2 \\ 3+i & -4i \end{bmatrix}$,

$$A^t = \begin{bmatrix} 1+i & 3-i \\ 2 & 4i \end{bmatrix} \quad \text{y} \quad A^* = \begin{bmatrix} 1-i & 3+i \\ 2 & -4i \end{bmatrix}.$$

EJEMPLO 5. Las dos matrices $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ y $\begin{bmatrix} 1 & 2+i \\ 2-i & 3 \end{bmatrix}$ son hermitianas.

La primera es simétrica, la segunda no.

EJEMPLO 6. Las dos matrices $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$ y $\begin{bmatrix} i & -2 \\ 2 & 3i \end{bmatrix}$ son hemi-hermitianas.

La primera es hemi-simétrica, la segunda no.

EJEMPLO 7. Todos los elementos diagonales de una matriz hermitiana son reales. Todos los elementos diagonales de una matriz hemi-hermitiana son imaginarios puros. Todos los elementos diagonales de una matriz hemi-simétrica son cero.

EJEMPLO 8. Para una matriz cualquiera A, la matriz $B = \frac{1}{2}(A + A^*)$ es hermitiana, y la matriz $C = \frac{1}{2}(A - A^*)$ es hemi-hermitiana. Su suma es A. Así pues, toda matriz cuadrada A puede expresarse como una suma $A = B + C$, sien-

do B hermitiana y C hemi-hermitiana. Es un ejercicio sencillo comprobar que esa descomposición es única. Asimismo toda matriz cuadrada A puede expresarse de manera única como suma de una matriz simétrica, $\frac{1}{2}(A + A^t)$, y una matriz hemi-simétrica, $\frac{1}{2}(A - A^t)$.

EJEMPLO 9. Si A es ortogonal tenemos $1 = \det(AA^t) = (\det A)(\det A^t) = (\det A)^2$, así que $\det A = \pm 1$.

5.11 Ejercicios

1. Determinar cuáles de las siguientes matrices son simétricas, hemi-simétricas, hermitianas o hemi-hermitianas.

$$\text{a)} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 4 \end{bmatrix}, \qquad \text{b)} \begin{bmatrix} 0 & i & 2 \\ i & 0 & 3 \\ -2 & -3 & 4i \end{bmatrix},$$

$$\text{c)} \begin{bmatrix} 0 & i & 2 \\ -i & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}, \qquad \text{d)} \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}.$$

2. a) Comprobar que la matriz 2×2 $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ es una matriz ortogonal.

   b) Sea T la transformación lineal con la matriz anterior A relativa a la base ordinaria $\{i, j\}$. Demostrar que T aplica cada punto del plano de coordenadas polares (r, α) en el punto $(r, \alpha + \theta)$. Así pues, T es una rotación del plano alrededor del origen, siendo θ el ángulo de giro.

3. Sea V el espacio real de dimensión 3 con la base i, j, k. Demostrar que cada una de las matrices siguientes es ortogonal y representa la transformación que se indica.

$$\text{a)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad \text{(simetría en el plano } x, y).$$

$$\text{b)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad \text{(simetría respecto al eje } x).$$

c) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ (simetría respecto al origen).

d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ (rotación alrededor del eje x).

e) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ (rotación alrededor del eje x seguida de simetría respecto al plano yz).

4. Una matriz ortogonal real A se llama *propia* si $\det A = 1$, e *impropia* si $\det A = -1$.

a) Si A es una matriz propia 2×2, demostrar que $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ para un cierto θ. Ésta representa una rotación de ángulo θ.

b) Demostrar que $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ y $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ son matrices impropias. La primera representa una simetría en el plano xy respecto al eje x; la segunda representa una simetría respecto al eje y. Hallar todas las matrices impropias 2×2.

En cada uno de los ejercicios del 5 al 8, hallar a) un conjunto ortogonal de autovectores para A, y b) una matriz unitaria C tal que $C^{-1}AC$ sea una matriz diagonal.

5. $A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$.

6. $A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$.

7. $A = \begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$.

8. $A = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$.

9. Determinar cuáles de las siguientes matrices son unitarias, y cuáles son ortogonales (a, b, θ reales).

a) $\begin{bmatrix} e^{ia} & 0 \\ 0 & e^{ib} \end{bmatrix}$, b) $\begin{bmatrix} \cos\theta & 0 & \operatorname{sen}\theta \\ 0 & 1 & 0 \\ -\operatorname{sen}\theta & 0 & \cos\theta \end{bmatrix}$, c) $\begin{bmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{3}\sqrt{3} & \frac{1}{6}\sqrt{6} \\ 0 & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} & \frac{1}{3}\sqrt{3} & -\frac{1}{6}\sqrt{6} \end{bmatrix}$.

10. La teoría especial de la relatividad hace uso de las ecuaciones

$$x' = a(x - vt), \qquad y' = y, \qquad z' = z, \qquad t' = a(t - vx/c^2).$$

Aquí v es la velocidad de un objeto móvil, c la velocidad de la luz, y $a = c/\sqrt{c^2 - v^2}$. La transformación lineal que aplica (x, y, z, t) en (x', y', z', t') se llama *transformación de Lorentz*. a) Pongamos $(x_1, x_2, x_3, x_4) = (x, y, z, ict)$ y $(x_1', x_2', x_3', x_4') = (x', y', z', ict')$. Demostrar que las cuatro ecuaciones pueden ponerse en forma de una ecuación matricial como sigue,

$$[x_1', x_2', x_3', x_4'] = [x_1, x_2, x_3, x_4]\begin{bmatrix} a & 0 & 0 & -iav/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ iav/c & 0 & 0 & a \end{bmatrix}.$$

b) Demostrar que la matriz 4×4 de a) es ortogonal pero no unitaria.

11. Sean a un número real no nulo y A la matriz hemi-simétrica $A = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$.

a) Hallar un conjunto ortonormal de autovectores para A.
b) Hallar una matriz unitaria C tal que $C^{-1}AC$ sea una matriz diagonal.
c) Demostrar que no existe una matriz ortogonal real C tal que $C^{-1}AC$ sea una matriz diagonal.

12. Si los autovalores de una matriz hermitiana o hemi-hermitiana A son todos iguales a c, demostrar que $A = cI$.
13. Si A es una matriz hemi-simétrica real, demostrar que las dos matrices $I - A$ e $I + A$ son no singulares y que $(I - A)(I + A)^{-1}$ es ortogonal.
14. Para cada una de las siguientes proposiciones relativas a las matrices $n\times n$, dar una demostración o un contraejemplo.

a) Si A y B son unitarias, entonces $A + B$ es unitaria.
b) Si A y B son unitarias, AB es unitaria.
c) Si A y AB son unitarias, B es unitaria.
d) Si A y B son unitarias, $A + B$ no es unitaria.

5.12 Formas cuadráticas

Sean V un espacio euclídeo *real* y $T: V \to V$ un operador simétrico. Esto significa que T puede cambiarse de un factor al otro en un producto interior

$$(T(x), y) = (x, T(y)) \quad \text{para todo par } x, y \text{ de } V.$$

Dada T, definamos una función de valores reales Q en V mediante la ecuación

$$Q(x) = (T(x), x).$$

La función Q se denomina *forma cuadrática* asociada a T. La palabra «cuadrática» está sugerida por el teorema siguiente que muestra que en el caso de dimensión finita $Q(x)$ es un polinomio cuadrático en los componentes de x.

TEOREMA 5.8. *Sea $(e_1, \ldots, e_n)$ una base ortonormal para un espacio euclídeo V. Sean $T: V \to V$ una transformación simétrica y $A = (a_{ij})$ la matriz de T relativa a esa base. Entonces la forma cuadrática $Q(x) = (T(x), x)$ está ligada a A del modo siguiente:*

$$(5.7) \qquad Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \qquad \text{si} \quad x = \sum_{i=1}^{n} x_i e_i .$$

Demostración. En virtud de la linealidad tenemos $T(x) = \sum x_i T(e_i)$. Por consiguiente,

$$Q(x) = \left(\sum_{i=1}^{n} x_i T(e_i), \sum_{j=1}^{n} x_j e_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j (T(e_i), e_j) .$$

Esto demuestra (5.7) ya que $a_{ij} = \bar{a}_{ji} = (T(e_i), e_j)$.

La suma que aparece en (5.7) tiene sentido incluso si la matriz A no es simétrica.

DEFINICIÓN. *Sea V un espacio euclídeo real cualquiera con una base ortonormal $(e_1, \ldots, e_n)$, y sea $A = (a_{ij})$ cualquier matriz $n \times n$ de escalares. La función de valores escalares Q definida en cada elemento $x = \sum x_i e_i$ de V por la suma doble*

$$(5.8) \qquad Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

se denomina la forma cuadrática asociada a A.

Si A es una matriz diagonal, entonces $a_{ij} = 0$ si $i \neq j$ de modo que la suma (5.8) contiene únicamente cuadrados y se puede poner en la forma más sencilla

$$(5.9) \qquad Q(x) = \sum_{i=1}^{n} a_{ii} x_i^2 .$$

En este caso la forma cuadrática se llama *forma diagonal.*

La suma doble que aparece en (5.8) también se puede expresar como un producto de tres matrices.

TEOREMA 5.9. *Sea $X = [x_1, \ldots, x_n]$ una matriz fila $1\times n$, y sea $A = (a_{ij})$ una matriz $n\times n$. Entonces XAX^t es una matriz 1×1 cuyo elemento es*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j .$$

Demostración. El producto XA es una matriz $1\times n$, $XA = [y_1, \ldots, y_n]$, en donde el elemento y_j es el producto escalar de X por la columna j de A,

$$y_j = \sum_{i=1}^{n} x_i a_{ij} .$$

Por consiguiente, el producto XAX^t es una matriz 1×1 cuyo único elemento es el producto escalar

$$\sum_{j=1}^{n} y_j x_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} x_i a_{ij}\right)x_j = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j .$$

Observación: Es corriente identificar la matriz, 1×1, XAX^t con la suma (5.9) y llamar al producto XAX^t forma cuadrática. La ecuación (5.8) se escribe más sencillamente así:

$$Q(x) = XAX^t .$$

EJEMPLO 1. Sean $A = \begin{bmatrix} 1 & -1 \\ -3 & 5 \end{bmatrix}$, $X = [x_1, x_2]$. Tenemos entonces

$$XA = [x_1, x_2]\begin{bmatrix} 1 & -1 \\ -3 & 5 \end{bmatrix} = [x_1 - 3x_2, -x_1 + 5x_2],$$

y por tanto

$$XAX^t = [x_1 - 3x_2, -x_1 + 5x_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 3x_2x_1 - x_1x_2 + 5x_2^2 .$$

EJEMPLO 2. Sean $B = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$, $X = [x_1, x_2]$. Tenemos entonces

$$XBX^t = [x_1, x_2]\begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 2x_2x_1 - 2x_1x_2 + 5x_2^2 .$$

En ambos ejemplos 1 y 2, los términos en los productos mixtos dan al sumarlos $-4x_1x_2$ así que $XAX^t = XBX^t$. Estos ejemplos ponen de manifiesto que matrices distintas pueden conducir a la misma forma cuadrática. Obsérvese que una de esas matrices es simétrica. Esto constituye un ejemplo del siguiente teorema.

TEOREMA 5.10. *Para cualquier matriz A, $n\times n$, y cualquier matriz fila X, $1\times n$, tenemos $XAX^t = XBX^t$ en donde B es la matriz simétrica $B = \frac{1}{2}(A + A^t)$.*

Demostración. Puesto que XAX^t es una matriz 1×1, es igual a su transpuesta, $XAX^t = (XAX^t)^t$. Pero la transpuesta de un producto es el producto de las transpuestas en orden inverso, por lo que $(XAX^t)^t = XA^tX^t$. Por consiguiente $XAX^t = \frac{1}{2}XAX^t + \frac{1}{2}XA^tX^t = XBX^t$.

5.13 Reducción de una forma cuadrática real a forma diagonal

Una matriz simétrica real A es hermitiana. Por tanto, según el teorema 5.7 es semejante a la matriz diagonal $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, de sus autovalores. Además, tenemos $\Lambda = C^tAC$, siendo C una matriz ortogonal. Vamos ahora a demostrar que C puede utilizarse para convertir la forma cuadrática XAX^t a una forma diagonal.

TEOREMA 5.11. *Sea XAX^t una forma cuadrática asociada a una matriz simétrica real A, y sea C una matriz ortogonal que convierte A en una matriz diagonal $\Lambda = C^tAC$. Tenemos entonces*

$$XAX^t = Y\Lambda Y^t = \sum_{i=1}^{n} \lambda_i y_i^2,$$

donde $Y = [y_1, \ldots, y_n]$ es la matriz fila $Y = XC$, y $\lambda_1, \ldots, \lambda_n$ son los autovalores de A.

Demostración. Puesto que C es ortogonal tenemos $C^{-1} = C^t$. Por consiguiente la ecuación $Y = XC$ implica $X = YC^t$, y obtenemos

$$XAX^t = (YC^t)A(YC^t)^t = Y(C^tAC)Y^t = Y\Lambda Y^t.$$

Observación: El teorema 5.11 se expresa diciendo que la transformación lineal $Y = XC$ reduce la forma cuadrática XAX^t a una forma diagonal $Y\Lambda Y^t$.

EJEMPLO 1. La forma cuadrática correspondiente a la matriz identidad es

$$XIX^t = \sum_{i=1}^{n} x_i^2 = \|X\|^2,$$

que es el cuadrado de la longitud del vector $X = (x_1, \ldots, x_n)$. Una transformación lineal $Y = XC$, donde C es una matriz ortogonal, nos da una nueva forma cuadrática $Y\Lambda Y^t$ siendo $\Lambda = CIC^t = CC^t = I$. Puesto que $XIX^t = YIY^t$ tenemos $\|X\|^2 = \|Y\|^2$, por lo que Y tiene la misma longitud que X. Una transformación lineal que conserva la longitud de cada vector se denomina *isometría*.

En la sección 5.19 se estudian con mayor detalle esas transformaciones.

EJEMPLO 2. Determinar una matriz ortogonal C que reduzca la forma cuadrática $Q(x) = 2x_1^2 + 4x_1x_2 + 5x_2^2$ a forma diagonal.

Solución. Escribamos $Q(x) = XAX^t$, siendo $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$. Esta matriz simétrica se diagonalizó en el ejemplo 1 que sigue al teorema 5.7. Tiene los autovalores $\lambda_1 = 1$, $\lambda_2 = 6$, y un conjunto ortonormal de autovectores u_1, u_2, siendo $u_1 = t(2, -1)$, $u_2 = t(1, 2)$, $t = 1/\sqrt{5}$. Una matriz diagonalizante ortogonal es $C = t\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. La forma diagonal correspondiente es

$$Y\Lambda Y^t = \lambda_1 y_1^2 + \lambda_2 y_2^2 = y_1^2 + 6y_2^2.$$

El resultado del ejemplo 2 tiene una interpretación geométrica sencilla, representada en la figura 5.1. La transformación lineal $Y = XC$ puede considerarse como una rotación que aplica la base $\boldsymbol{i}, \boldsymbol{j}$ en la nueva base u_1, u_2. Un punto que respecto a la primera base tiene coordenadas (x_1, x_2) tiene las nuevas coordenadas (y_1, y_2) respecto a la segunda base. Puesto que $XAX^t = Y\Lambda Y^t$, el conjunto de puntos (x_1, x_2) que satisfacen la ecuación $XAX^t = c$ para un cierto c es idéntico al conjunto de puntos (y_1, y_2) que satisfacen $Y\Lambda Y^t = c$. La segunda ecuación, escrita en la forma $y_1^2 + 6y_2^2 = c$, es la ecuación cartesiana de una elipse si $c > 0$. Por tanto la ecuación $XAX^t = c$, puesta en la forma $2x_1^2 + 4x_1x_2 + 5x_2^2 = c$, representa la misma elipse en el sistema coordenado original. La figura 5.1 muestra la elipse correspondiente a $c = 9$.

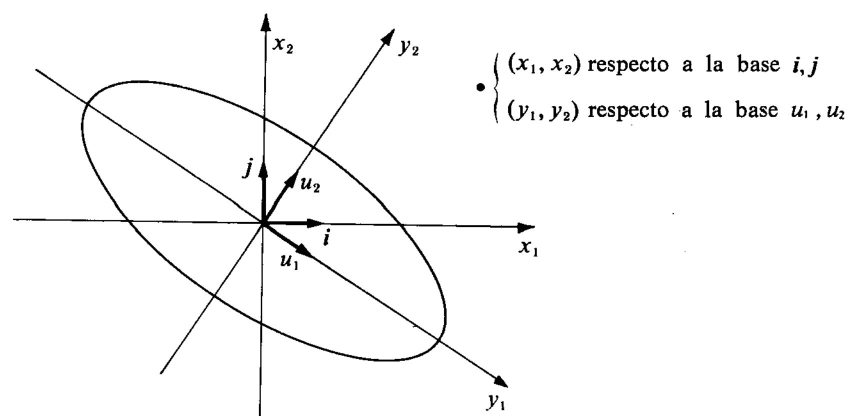


FIGURA 5.1 *Rotación de ejes mediante una matriz ortogonal. La elipse tiene la ecuación cartesiana* $XAX^t = 9$ *en el sistema* x_1x_2, *y la ecuación* $Y\Lambda Y^t = 9$ *en el sistema* y_1y_2.

5.14 Aplicaciones a la Geometría Analítica

La reducción de una forma cuadrática a forma diagonal puede utilizarse para identificar el conjunto de todos los puntos (x, y) del plano que satisface una ecuación cartesiana de la forma

$$(5.10) \qquad ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

Encontraremos siempre que ese conjunto es una sección cónica, esto es, una elipse, una hipérbola, una parábola o uno de los casos degenerados (el conjunto vacío, un solo punto, o una o dos rectas). El tipo de cónica viene caracterizado por los términos de segundo grado, esto es, por la forma cuadrática $ax^2 + bxy + cy^2$. Téngase en cuenta que aquí hemos puesto x en lugar de x_1 e y en lugar de x_2. Volviendo a la notación anterior y expresando esa forma cuadrática como producto matricial

$$XAX^t = ax_1^2 + bx_1x_2 + cx_2^2,$$

siendo $X = [x_1, x_2]$ y $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$. Con la rotación $Y = XC$ reducimos esa forma cuadrática a una forma diagonal $\lambda_1 y_1^2 + \lambda_2 y_2^2$, donde λ_1, λ_2 son los autovalores de A. Un conjunto ortonormal de autovectores u_1, u_2 determina un nuevo conjunto de ejes coordenados, respecto a los cuales la ecuación cartesiana (5.10) se convierte en

$$(5.11) \qquad \lambda_1 y_1^2 + \lambda_2 y_2^2 + d'y_1 + e'y_2 + f = 0,$$

con nuevos coeficientes d' y e' en los términos de primer grado. En esta ecuación no existe término en el producto mixto y_1y_2, con lo cual el tipo de cónica se identifica fácilmente examinando los autovalores λ_1 y λ_2. Si la cónica no es degenerada, la ecuación (5.11) representa una *elipse* si λ_1, λ_2 tienen el mismo signo, una *hipérbola* si λ_1, λ_2 tienen signos contrarios, y una *parábola* si uno de los dos es cero y el otro no. Los tres casos corresponden a $\lambda_1\lambda_2 > 0$, $\lambda_1\lambda_2 < 0$, y $\lambda_1\lambda_2 = 0$. Seguidamente damos algunos ejemplos.

EJEMPLO 1. $2x^2 + 4xy + 5y^2 + 4x + 13y - \frac{1}{4} = 0$. Escribimos la ecuación así

$$(5.12) \qquad 2x_1^2 + 4x_1x_2 + 5x_2^2 + 4x_1 + 13x_2 - \tfrac{1}{4} = 0.$$

La forma cuadrática $2x_1^2 + 4x_1x_2 + 5x_2^2$ es la tratada en el ejemplo 2 de la sección

anterior. Su matriz tiene los autovalores $\lambda_1 = 1$, $\lambda_2 = 6$, y un conjunto ortonormal de autovectores $u_1 = t(2, -1)$, $u_2 = t(1, 2)$, siendo $t = 1/\sqrt{5}$. Una matriz diagonalizante ortogonal es $C = t\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. Ésta reduce la parte cuadrática de (5.12) a la forma $y_1^2 + 6y_2^2$. Para determinar el efecto en la parte de primer grado ponemos la ecuación de la rotación $Y = XC$ en la forma $X = YC^t$ y obtenemos

$$[x_1, x_2] = \frac{1}{\sqrt{5}}[y_1, y_2]\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix},$$

$$x_1 = \frac{1}{\sqrt{5}}(2y_1 + y_2), \qquad x_2 = \frac{1}{\sqrt{5}}(-y_1 + 2y_2).$$

Por consiguiente, la parte lineal $4x_1 + 13x_2$ se transforma en

$$\frac{4}{\sqrt{5}}(2y_1 + y_2) + \frac{13}{\sqrt{5}}(-y_1 + 2y_2) = -\sqrt{5}\, y_1 + 6\sqrt{5}\, y_2.$$

La ecuación cartesiana transformada se convierte en

$$y_1^2 + 6y_2^2 - \sqrt{5}\, y_1 + 6\sqrt{5}\, y_2 - \tfrac{1}{4} = 0.$$

Completando los cuadrados en y_1 e y_2 escribimos ésa así:

$$(y_1 - \tfrac{1}{2}\sqrt{5})^2 + 6(y_2 + \tfrac{1}{2}\sqrt{5})^2 = 9.$$

Ésta es la ecuación de una elipse con su centro en el punto $(\frac{1}{2}\sqrt{5}, -\frac{1}{2}\sqrt{5})$ en el sistema coordenado y_1y_2. Las direcciones positivas de los ejes y_1 e y_2 están determinadas por los autovectores u_1 y u_2, como se indica en la figura 5.2.

Podemos aún simplificar más la ecuación escribiendo

$$z_1 = y_1 - \tfrac{1}{2}\sqrt{5}, \qquad z_2 = y_2 + \tfrac{1}{2}\sqrt{5}.$$

Geométricamente, esto es lo mismo que introducir un nuevo sistema de ejes coordenados paralelos a los ejes y_1y_2 pero con el nuevo origen en el centro de la elipse. En el sistema z_1z_2 la ecuación de la elipse es sencilla

$$z_1^2 + 6z_2^2 = 9, \qquad \text{o} \qquad \frac{z_1^2}{9} + \frac{z_2^2}{3/2} = 1.$$

La elipse y los tres sistemas de coordenadas se han dibujado en la figura 5.2.

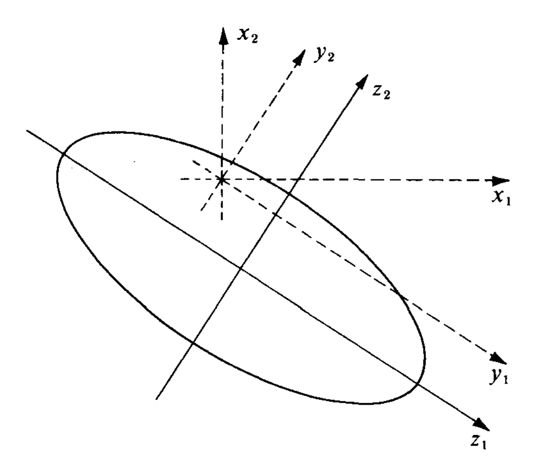


FIGURA 5.2 *Rotación y traslación de ejes coordenados. La rotación $Y = XC$ viene seguida por la traslación $z_1 = y_1 - \frac{1}{2}\sqrt{5}$, $z_2 = y_2 + \frac{1}{2}\sqrt{5}$.*

EJEMPLO 2. $2x^2 - 4xy - y^2 - 4x + 10y - 13 = 0$. Pongámosla en la forma

$$2x_1^2 - 4x_1x_2 - x_2^2 - 4x_1 + 10x_2 - 13 = 0.$$

La parte cuadrática es XAX^t, siendo $A = \begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix}$. Esta matriz tiene los autovalores $\lambda_1 = 3$, $\lambda_2 = -2$. Un conjunto ortonormal de autovectores es $u_1 = t(2, -1)$, $u_2 = t(1, 2)$, siendo $t = 1/\sqrt{5}$. Una matriz diagonalizante ortogonal es $C = t\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. La ecuación de la rotación $X = YC^t$ nos da

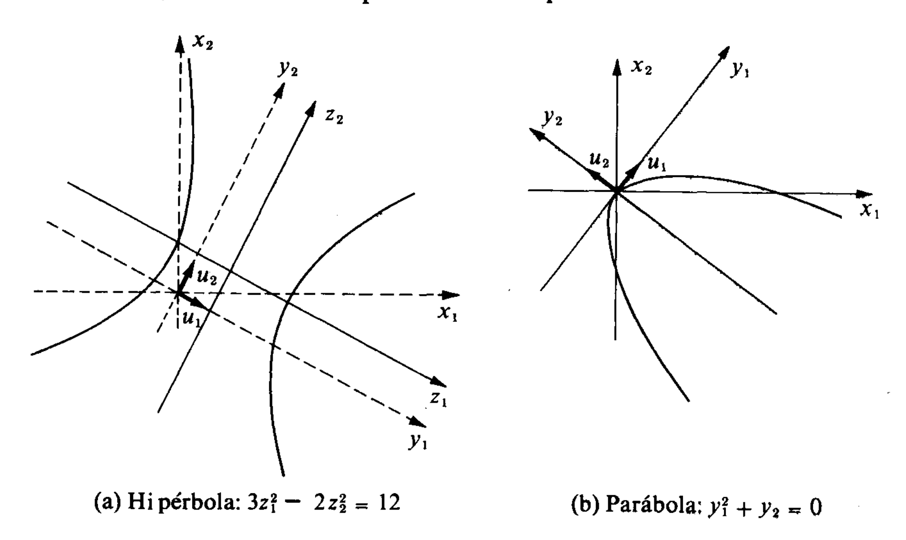
$$x_1 = \frac{1}{\sqrt{5}}(2y_1 + y_2), \qquad x_2 = \frac{1}{\sqrt{5}}(-y_1 + 2y_2).$$

Por consiguiente, la ecuación transformada se convierte en

$$3y_1^2 - 2y_2^2 - \frac{4}{\sqrt{5}}(2y_1 + y_2) + \frac{10}{\sqrt{5}}(-y_1 + 2y_2) - 13 = 0,$$

o

$$3y_1^2 - 2y_2^2 - \frac{18}{\sqrt{5}}y_1 + \frac{16}{\sqrt{5}}y_2 - 13 = 0.$$



(a) Hipérbola: $3z_1^2 - 2z_2^2 = 12$ (b) Parábola: $y_1^2 + y_2 = 0$

FIGURA 5.3 *Curvas de los ejemplos 2 y 3.*

Completando los cuadrados en y_1 e y_2 obtenemos la ecuación

$$3(y_1 - \tfrac{3}{5}\sqrt{5})^2 - 2(y_2 - \tfrac{4}{5}\sqrt{5})^2 = 12,$$

que representa una hipérbola con centro en $(\frac{3}{5}\sqrt{5}, \frac{4}{5}\sqrt{5})$ en el sistema y_1y_2. La traslación $z_1 = y_1 - \frac{3}{5}\sqrt{5}$, $z_2 = y_2 - \frac{4}{5}\sqrt{5}$ simplifica la ecuación llegando a

$$3z_1^2 - 2z_2^2 = 12, \qquad \text{o} \qquad \frac{z_1^2}{4} - \frac{z_2^2}{6} = 1.$$

La hipérbola está dibujada en la figura 5.3(a). Los autovectores u_1 y u_2 determinan las direcciones positivas de los ejes y_1 e y_2.

EJEMPLO 3. $9x^2 + 24xy + 16y^2 - 20x + 15y = 0$. Pongamos esta ecuación así:

$$9x_1^2 + 24x_1x_2 + 16x_2^2 - 20x_1 + 15x_2 = 0.$$

La matriz simétrica para la parte cuadrática es $A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$. Sus autovalores son $\lambda_1 = 25$, $\lambda_2 = 0$. Un conjunto ortonormal de autovectores es $u_1 = \frac{1}{5}(3, 4)$,

$u_2 = \frac{1}{5}(-4, 3)$. Una matriz diagonalizante ortogonal es $C = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$. La rotación $X = YC^t$ nos da

$$x_1 = \tfrac{1}{5}(3y_1 - 4y_2), \qquad x_2 = \tfrac{1}{5}(4y_1 + 3y_2).$$

Por consiguiente, la ecuación cartesiana transformada se convierte en

$$25y_1^2 - \tfrac{20}{5}(3y_1 - 4y_2) + \tfrac{15}{5}(4y_1 + 3y_2) = 0.$$

Simplificando esa ecuación se llega a $y_1^2 + y_2 = 0$, que es la ecuación de una parábola con el vértice en el origen. La parábola está dibujada en la figura 5.3(b).

EJEMPLO 4. *Casos degenerados.* El solo conocimiento de los autovalores no revela si la ecuación cartesiana representa una cónica degenerada. Por ejemplo, las tres ecuaciones $x^2 + 2y^2 = 1$, $x^2 + 2y^2 = 0$ y $x^2 + 2y^2 = -1$ tienen los mismos autovalores; la primera representa una elipse no degenerada, la segunda sólo se satisface para $(x, y) = (0, 0)$, y la tercera representa el conjunto vacío. Las dos últimas pueden considerarse casos degenerados de elipse.

La gráfica de la ecuación $y^2 = 0$ es el eje x. La ecuación $y^2 - 1 = 0$ representa las dos rectas paralelas $y = 1$ e $y = -1$. Éstos pueden considerarse como casos degenerados de parábola. La ecuación $x^2 - 4y^2 = 0$ representa dos rectas que se cortan ya que se satisface si $x - 2y = 0$ o si $x + 2y = 0$. Éste puede considerarse como caso degenerado de hipérbola.

Sin embargo, si la ecuación cartesiana $ax^2 + bxy + cy^2 + dx + ey + f = 0$ representa una cónica no degenerada, entonces el *tipo* de cónica puede determinarse con toda facilidad. El polinomio característico de la matriz de la forma cuadrática $ax^2 + bxy + cy^2$ es

$$\det \begin{bmatrix} \lambda - a & -b/2 \\ -b/2 & \lambda - c \end{bmatrix} = \lambda^2 - (a + c)\lambda + (ac - \tfrac{1}{4}b^2) = (\lambda - \lambda_1)(\lambda - \lambda_2).$$

Por tanto, el producto de los autovalores es

$$\lambda_1\lambda_2 = ac - \tfrac{1}{4}b^2 = \tfrac{1}{4}(4ac - b^2).$$

Puesto que el tipo de cónica está determinado por el signo algebraico del producto $\lambda_1\lambda_2$, vemos que la cónica es *elipse, hipérbola* o *parábola,* según que $4ac - b^2$ sea *positivo, negativo* o *nulo.* El número $4ac - b^2$ se llama *discriminante* de la forma cuadrática $ax^2 + bxy + cy^2$. En los ejemplos 1, 2 y 3 el discriminante tiene los valores 34, -24, y 0, respectivamente.

5.15 Ejercicios

En cada uno de los ejercicios 1 al 7, hallar a) una matriz simétrica A para la forma cuadrática; b) los autovalores de A; c) un conjunto ortonormal de autovectores; d) una matriz diagonalizante ortogonal C.

1. $4x_1^2 + 4x_1x_2 + x_2^2$.
2. x_1x_2.
3. $x_1^2 + 2x_1x_2 - x_2^2$.
4. $34x_1^2 - 24x_1x_2 + 41x_2^2$.
5. $x_1^2 + x_1x_2 + x_1x_3 + x_2x_3$.
6. $2x_1^2 + 4x_1x_3 + x_2^2 - x_3^2$.
7. $3x_1^2 + 4x_1x_2 + 8x_1x_3 + 4x_2x_3 + 3x_3^2$.

En cada uno de los ejercicios 8 al 18, identificar la cónica representada por la ecuación cartesiana y hacer un dibujo.

8. $y^2 - 2xy + 2x^2 - 5 = 0$.
9. $y^2 - 2xy + 5x = 0$.
10. $y^2 - 2xy + x^2 - 5x = 0$.
11. $5x^2 - 4xy + 2y^2 - 6 = 0$.
12. $19x^2 + 4xy + 16y^2 - 212x + 104y = 356$.
13. $9x^2 + 24xy + 16y^2 - 52x + 14y = 6$.
14. $5x^2 + 6xy + 5y^2 - 2 = 0$.
15. $x^2 + 2xy + y^2 - 2x + 2y + 3 = 0$.
16. $2x^2 + 4xy + 5y^2 - 2x - y - 4 = 0$.
17. $x^2 + 4xy - 2y^2 - 12 = 0$.
18. $xy + y - 2x - 2 = 0$.

19. ¿Para qué valor (o valores) de c la gráfica de la ecuación cartesiana $2xy - 4x + 7y + c = 0$ será un par de rectas?
20. Si la ecuación $ax^2 + bxy + cy^2 = 1$ representa una elipse, demostrar que el área de la región que limita es $2\pi/\sqrt{4ac - b^2}$. Esto da un significado geométrico al discriminante $4ac - b^2$.

* 5.16 (*) Autovalores de una transformación simétrica obtenidos como valores de su forma cuadrática

Vamos a prescindir ahora de la exigencia de que V sea de dimensión finita y vamos a encontrar una relación entre los autovalores de un operador simétrico y su forma cuadrática.

Supongamos que x es un autovector con norma 1 perteneciente a un autovalor λ. Entonces $T(x) = \lambda x$ con lo que tenemos

(5.13)
$$Q(x) = (T(x), x) = (\lambda x, x) = \lambda(x, x) = \lambda,$$

ya que $(x, x) = 1$. El conjunto de todos los x de V que satisfacen $(x, x) = 1$ se

(*) Las secciones señaladas con asterisco pueden suprimirse o posponerse sin perjuicio de la continuidad del tema.

llama la *esfera unidad* de V. La ecuación (5.13) demuestra el teorema siguiente.

TEOREMA 5.12. *Sea $T: V \to V$ una transformación simétrica en un espacio euclídeo real V, y sea $Q(x) = (T(x), x)$. Entonces los autovalores de T (si existe alguno) se encuentran entre los valores que Q toma en la esfera unidad de V.*

EJEMPLO. Sea $V = V_2(\mathbf{R})$ con la base $(\boldsymbol{i}, \boldsymbol{j})$ y con el producto escalar ordinario como producto interior. Sea T la transformación simétrica con matriz $A = \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}$. Entonces la forma cuadrática de T viene dada por

$$Q(x) = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij} x_i x_j = 4x_1^2 + 8x_2^2 .$$

Los autovalores de T son $\lambda_1 = 4$, $\lambda_2 = 8$. Es fácil ver que esos autovalores son, respectivamente, los valores mínimo y máximo que toma Q en el círculo unidad $x_1^2 + x_2^2 = 1$. En efecto, en ese círculo tenemos

$$Q(x) = 4(x_1^2 + x_2^2) + 4x_2^2 = 4 + 4x_2^2, \qquad \text{donde} \quad -1 \leq x_2 \leq 1 .$$

Éste alcanza el valor mínimo, 4, cuando $x_2 = 0$ y el máximo, 8, cuando $x_2 = \pm 1$.

La figura 5.4 muestra el círculo unidad y dos elipses. La elipse interior tiene como ecuación cartesiana $4x_1^2 + 8x_2^2 = 4$. Consta de todos los puntos $x = (x_1, x_2)$ del plano que satisfacen $Q(x) = 4$. La elipse exterior es la de ecuación $4x_1^2 + 8x_2^2 = 8$ y consta de los puntos que satisfacen $Q(x) = 8$. Los puntos $(\pm 1, 0)$ en los que la elipse interior es tangente al círculo unidad son autovectores correspondientes al autovalor 4. Los puntos $(0, \pm 1)$ de la elipse exterior son autovectores correspondientes al autovalor 8.

El ejemplo anterior ilustra propiedades relativas a los extremos de los autovalores que son válidos con mayor generalidad. En la sección que sigue probaremos que el mínimo y el máximo autovalor (si existen) son siempre el mínimo y el máximo valor que Q toma en la esfera unidad. En la discusión que vamos a hacer de esas propiedades utilizaremos el siguiente teorema relativo a formas cuadráticas. Debe observarse que este teorema no exige que V sea de dimensión finita.

TEOREMA 5.13. *Sea $T: V \to V$ una transformación simétrica en un espacio euclídeo real V con forma cuadrática $Q(x) = (T(x), x)$. Supongamos que Q no cambia de signo en V. Entonces si $Q(x) = 0$ para un cierto x de V también tenemos $T(x) = O$. Dicho de otro modo, si Q no cambia de signo, entonces Q se anula solamente en el núcleo de T.*

Demostración. Supongamos que $Q(x) = 0$ para un cierto x de V y sea y

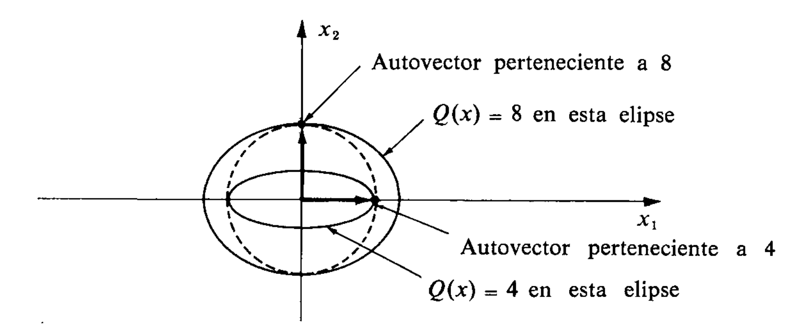


FIGURA 5.4 *Relación geométrica entre los autovalores de T y los valores de Q en la esfera unidad, ilustrada con un ejemplo bi-dimensional.*

un elemento cualquiera de V. Elijamos un t cualquiera y consideremos $Q(x + ty)$. Empleando la linealidad de T, la linealidad del producto interior, y la simetría de T, tenemos

$$\begin{aligned} Q(x + ty) &= (T(x + ty), x + ty) = (T(x) + tT(y), x + ty) \\ &= (T(x), x) + t(T(x), y) + t(T(y), x) + t^2(T(y), y) \\ &= Q(x) + 2t(T(x), y) + t^2Q(y) = at + bt^2, \end{aligned}$$

siendo $a = 2(T(x), y)$ y $b = Q(y)$. Si Q es no negativa en V tenemos la desigualdad

$$at + bt^2 \geq 0 \quad \text{para todo número real } t.$$

En otras palabras, el polinomio cuadrático $p(t) = at + bt^2$ tiene su mínimo en $t = 0$. Luego $p'(0) = 0$. Pero $p'(0) = a = 2(T(x), y)$, así que $(T(x), y) = 0$. Puesto que y es arbitrario, podemos tomar en particular $y = T(x)$, dando $(T(x), T(x)) = 0$. Esto demuestra que $T(x) = O$.

Si Q es no positiva en V tenemos $p(t) = at + bt^2 \leq 0$ para todo t, así que p tiene su máximo en $t = 0$, y por tanto $p'(0) = 0$ como antes.

* 5.17 Propiedades relativas a extremos de los autovalores de una transformación simétrica

Demostraremos ahora que los valores extremos de una forma cuadrática en la esfera unidad son autovalores.

TEOREMA 5.14. *Sea T: $V \to V$ una transformación lineal simétrica en un es-*

pacio euclídeo real V, y sea $Q(x) = (T(x), x)$. Entre todos los valores que Q toma en la esfera unidad, supongamos que existe un extremo (*) (máximo o mínimo) en un punto u en el que $(u, u) = 1$. Entonces u es un autovector para T; el correspondiente autovalor es $Q(u)$, valor extremo de Q en la esfera unidad.

Demostración. Supongamos que Q tiene un mínimo en u. Tenemos entonces

$$(5.14) \qquad Q(x) \geq Q(u) \quad \text{para todo } x \text{ que cumpla } (x, x) = 1.$$

Sea $\lambda = Q(u)$. Si $(x, x) = 1$ tenemos $Q(u) = \lambda(x, x) = (\lambda x, x)$ así que la desigualdad (5.14) puede ponerse así

$$(5.15) \qquad (T(x), x) \geq (\lambda x, x)$$

con tal que $(x, x) = 1$. Demostremos ahora que (5.15) es válido para *todo* x de V. Supongamos que $\|x\| = a$. Entonces $x = ay$, siendo $\|y\| = 1$. Luego

$$(T(x), x) = (T(ay), ay) = a^2(T(y), y) \qquad \text{y} \qquad (\lambda x, x) = a^2(\lambda y, y).$$

Pero $(T(y), y) \geq (\lambda y, y)$ ya que $(y, y) = 1$. Multiplicando ambos miembros de esta desigualdad por a^2 llegamos a (5.15) para $x = ay$.

Puesto que $(T(x), x) - (\lambda x, x) = (T(x) - \lambda x, x)$, podemos poner la desigualdad (5.15) en la forma $(T(x) - \lambda x, x) \geq 0$, o

$$(5.16) \qquad (S(x), x) \geq 0, \qquad \text{donde} \quad S = T - \lambda I.$$

Cuando $x = u$ tenemos la desigualdad (5.14) y por tanto también (5.16). La transformación lineal S es simétrica. La desigualdad (5.16) establece que la forma cuadrática Q_1 dada por $Q_1(x) = (S(x), x)$ es no negativa en V. Cuando $x = u$ tenemos $Q_1(x) = 0$. Por consiguiente, según el teorema 5.13 debe ser $S(u) = O$. Dicho de otro omdo, $T(u) = \lambda u$, así que u es un autovector para T, y $\lambda = Q(u)$ es el correspondiente autovalor. Esto completa la demostración si Q tiene un mínimo en u.

Si existe un máximo en u todas las desigualdades de la demostración anterior se invierten y aplicamos el teorema 5.13 a la forma cuadrática no positiva Q_1.

(*) Si V es de dimensión infinita, la forma cuadrática Q no tiene necesariamente un extremo en la esfera unidad. Este será el caso cuando T no tiene autovalores. En el caso de dimensión finita, Q tiene siempre un máximo y un mínimo en algún punto de la esfera unidad. Esto resulta como una consecuencia de un teorema más general sobre los valores extremos de funciones continuas. En la sección 9.16 puede verse un caso particular de ese teorema.

* 5.18 Caso de dimensión finita

Supongamos ahora que $\dim V = n$. Entonces T tiene n autovalores reales que pueden disponerse en orden creciente

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n .$$

De acuerdo con el teorema 5.14, el menor autovalor λ_1 es el mínimo de Q en la esfera unidad, y el mayor autovalor es el máximo de Q en la esfera unidad. Demostraremos seguidamente que los autovalores intermedios también se presentan como valores extremos de Q, restringidos a ciertos subconjuntos de la esfera unidad.

Sea u_1 un autovector en la esfera unidad que minimiza Q. Entonces $\lambda_1 = Q(u_1)$. Si λ es un autovalor distinto de λ_1 cualquier autovector correspondiente a λ debe ser ortogonal a u_1. Por consiguiente, es natural buscar como tal un autovector en el complemento ortogonal del subespacio engendrado por u_1.

Sea S el subespacio engendrado por u_1. El complemento ortogonal $S^\perp$ consta de todos los elementos de V ortogonal a u_1. En particular, $S^\perp$ contiene todos los autovectores correspondientes a los autovalores $\lambda \neq \lambda_1$. Es fácil comprobar que $\dim S^\perp = n - 1$ y que T aplica $S^\perp$ en sí mismo (*). Designemos por S_{n-1} la esfera unidad en el subespacio $S^\perp$ de dimensión $n - 1$. (La esfera unidad S_{n-1} es un subconjunto de la esfera unidad en V.) Aplicando el teorema 5.14 al subespacio $S^\perp$ encontramos que $\lambda_2 = Q(u_2)$, donde u_2 es un punto que hace mínimo Q en S_{n-1}.

El siguiente autovector λ_3 puede obtenerse en forma parecida como el valor mínimo de Q en la esfera unidad S_{n-2} en el espacio de dimensión $n - 2$ que consta de aquellos elementos ortogonales simultáneamente a u_1 y u_2. Continuando en esta forma encontramos que cada autovalor λ_k es el valor mínimo que toma Q en una esfera unidad S_{n-k+1} en un subespacio de dimensión $n - k + 1$. El mayor de esos mínimos, λ_n, es también el *máximo* valor que toma Q en cada una de las esferas S_{n-k+1}. El correspondiente conjunto de autovectores $u_1, \ldots, u_n$ forma una base ortonormal para V.

5.19 Transformaciones unitarias

Concluimos este capítulo con una breve discusión de otro tipo importante de transformaciones llamadas unitarias. En el caso de dimensión finita están representadas por matrices unidad.

DEFINICIÓN. *Sean E un espacio euclídeo y V un subespacio de E. Una transformación $T: V \to E$ se llama unitaria en V si tenemos*

$$(T(x), T(y)) = (x, y) \qquad \textit{para todo par } x, y \textit{ de } V. \tag{5.17}$$

(*) Esto se hizo en la demostración del teorema 5.4, sección 5.6.

Cuando E es un espacio euclídeo real, una transformación unitaria también se denomina transformación ortogonal.

La ecuación (5.17) se expresa diciendo que T conserva los productos interiores. En consecuencia, es natural esperar que T también conserve la ortogonalidad y las normas, pues éstas se derivan de los productos interiores.

TEOREMA 5.15. *Si $T: V \to E$ es una transformación unitaria de V, entonces para todo par x, y de V tenemos*

a) $(x, y) = 0$ *implica* $(T(x), T(y)) = 0$ (*T conserva la ortogonalidad*).
b) $\|T(x)\| = \|x\|$ (*T conserva las normas*).
c) $\|T(x) - T(y)\| = \|x - y\|$ (*T conserva las distancias*).
d) *T es invertible, y T^{-1} es unitaria en $T(V)$.*

Demostración. La parte a) es consecuencia inmediata de la ecuación (5.17). La parte b) se deduce haciendo $x = y$ en (5.17). La parte c) resulta de b) porque $T(x) - T(y) = T(x - y)$.

Para demostrar d) aplicamos b) lo que hace ver que $T(x)=O$ implica $x=O$, así que T es invertible. Si $x \in T(V)$ e $y \in T(V)$ podemos escribir $x = T(u)$, $y = T(v)$. con lo que tenemos

$$(T^{-1}(x), T^{-1}(y)) = (u, v) = (T(u), T(v)) = (x, y).$$

Por lo tanto T^{-1} es unitaria en $T(V)$.

En cuanto a los autovalores y autovectores tenemos el teorema siguiente.

TEOREMA 5.16. *Sea $T: V \to E$ una transformación unitaria en V.*

a) *Si T tiene un autovalor λ, entonces $|\lambda| = 1$.*
b) *Si x e y son autovectores correspondientes a autovalores distintos λ y μ, entonces x e y son ortogonales.*
c) *Si $V = E$ y $\dim V = n$, y si V es un espacio complejo, existen los autovectores $u_1, \ldots, u_n$ de T que forman una base ortonormal para V. La matriz de T relativa a esa base es la matriz diagonal $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, siendo λ_k el autovalor correspondiente a u_k.*

Demostración. Para demostrar a), sea x un autovector correspondiente a λ. Entonces $x \neq O$ y $T(x) = \lambda x$. Haciendo $y = x$ en la ecuación (5.17) llegamos a

$$(\lambda x, \lambda x) = (x, x) \qquad \text{o} \qquad \lambda\bar{\lambda}(x, x) = (x, x).$$

Puesto que $(x, x) > 0$ y $\lambda\bar{\lambda} = |\lambda|^2$, esto implica $|\lambda| = 1$.

Para demostrar b), escribamos $T(x) = \lambda x$, $T(y) = \mu y$ y calcularemos el producto interior $(T(x), T(y))$ de dos maneras. Tenemos

$$(T(x), T(y)) = (x, y)$$

ya que T es unitaria. Asimismo

$$(T(x), T(y)) = (\lambda x, \mu y) = \lambda\bar{\mu}(x, y)$$

ya que x e y son autovectores. Por consiguiente $\lambda\bar{\mu}(x, y) = (x, y)$, así que $(x, y) = 0$ salvo si $\lambda\bar{\mu} = 1$. Pero $\lambda\bar{\lambda} = 1$ según a), por lo que si fuera $\lambda\bar{\mu} = 1$ debería ser $\lambda\bar{\lambda} = \lambda\bar{\mu}$, $\bar{\lambda} = \bar{\mu}$, $\lambda = \mu$, lo que contradice la hipótesis de que λ y μ son distintos. Por consiguiente $\lambda\bar{\mu} \neq 1$ y $(x, y) = 0$.

El apartado c) se demuestra por inducción respecto de n en forma muy parecida al teorema 5.4, que da el resultado análogo para los operadores hermitianos. El único cambio necesario afecta a la parte de la demostración que afirma que T aplica $S^{\perp}$ en sí mismo, donde

$$S^{\perp} = \{x \mid x \in V, \quad (x, u_1) = 0\}.$$

Aquí u_1 es un autovector de T con el autovalor λ_1. A partir de la ecuación $T(u_1) = \lambda_1 u_1$ encontramos

$$u_1 = \lambda_1^{-1}T(u_1) = \bar{\lambda}_1 T(u_1)$$

ya que $\lambda_1\bar{\lambda}_1 = |\lambda_1|^2 = 1$. Elijamos ahora un x cualquiera en $S^{\perp}$ y observemos que

$$(T(x), u_1) = (T(x), \bar{\lambda}_1 T(u_1)) = \lambda_1(T(x), T(u_1)) = \lambda_1(x, u_1) = 0.$$

Luego $T(x) \in S^{\perp}$ si $x \in S^{\perp}$, por lo cual T aplica $S^{\perp}$ en sí mismo. El resto de la demostración es idéntico a la del teorema 5.4, por lo que no lo repetiremos.

Los dos teoremas que siguen se refieren a propiedades de las transformaciones unitarias en espacios de dimensión finita. Damos tan sólo un bosquejo de las demostraciones.

TEOREMA 5.17. *Supongamos que* $\dim V = n$ *y sea* $E = (e_1, \ldots, e_n)$ *una base para* V *fijada. La transformación lineal* $T: V \to V$ *es unitaria si y sólo si*

$$(5.18) \qquad (T(e_i), T(e_j)) = (e_i, e_j) \quad \textit{para todo par } i, j.$$

En particular, si E es ortonormal entonces T es unitaria si y sólo si T aplica E en una base ortonormal.

Esbozo de la demostración. Pongamos $x = \sum x_i e_i$, $y = \sum y_j e_j$. Tenemos entonces

$$(x, y) = \left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{y}_j (e_i, e_j),$$

y

$$(T(x), T(y)) = \left(\sum_{i=1}^{n} x_i T(e_i), \sum_{j=1}^{n} y_j T(e_j)\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{y}_j (T(e_i), T(e_j)).$$

A continuación se compara (x, y) con $(T(x), T(y))$.

TEOREMA 5.18. *Supongamos que* $\dim V = n$ y sea $(e_1, \ldots, e_n)$ *una base ortonormal para* V. *Sea* $A = (a_{ij})$ *la representación matricial de una transformación lineal* $T: V \to V$ *relativa a esa base. Entonces T es unitaria si y sólo si A es unitaria, esto es, si y sólo si*

(5.19) $$A^*A = I.$$

Esbozo de la demostración. Puesto que (e_i, e_j) es el elemento ij de la matriz identidad, la ecuación (5.19) implica

(5.20) $$(e_i, e_j) = \sum_{k=1}^{n} \bar{a}_{ki} a_{kj} = \sum_{k=1}^{n} a_{ki} \bar{a}_{kj}.$$

Puesto que A es la matriz de T tenemos $T(e_i) = \sum_{k=1}^{n} a_{ki} e_k$, $T(e_j) = \sum_{r=1}^{n} a_{rj} e_r$ con lo cual

$$(T(e_i), T(e_j)) = \left(\sum_{k=1}^{n} a_{ki} e_k, \sum_{r=1}^{n} a_{rj} e_r\right) = \sum_{k=1}^{n} \sum_{r=1}^{n} a_{ki} \bar{a}_{rj} (e_k, e_r) = \sum_{k=1}^{n} a_{ki} \bar{a}_{kj}.$$

Se compara luego con (5.20) y se aplica el teorema 5.17.

TEOREMA 5.19. *Toda matriz unitaria A tiene las propiedades siguientes:*

a) *A es no singular y* $A^{-1} = A^*$.

b) *Cada una de las matrices* A^t, $\bar{A}$ y A^* *es unitaria.*

c) *Los autovalores de A son números complejos de valor absoluto* 1.

d) $|\det A| = 1$; *si A es real, entonces* $\det A = \pm 1$.

La demostración del teorema 5.19 se deja como ejercicio para el lector.

5.20 Ejercicios

1. a) Sea $T: V \to V$ la transformación dada por $T(x) = cx$, donde c es un escalar fijado. Demostrar que T es unitaria si y sólo si $|c| = 1$.
   b) Si V es unidimensional, demostrar que las únicas transformaciones unitarias en V son las descritas en a). En particular, si V es un espacio real unidimensional, existen sólo dos transformaciones ortogonales, $T(x) = x$ y $T(x) = -x$.
2. Demostrar cada una de las proposiciones relativas a una matriz A, $n \times n$, real ortogonal
   a) Si λ es un autovalor real de A, entonces $\lambda = 1$ ó $\lambda = -1$.
   b) Si λ es un autovalor complejo de A, el complejo conjugado $\bar{\lambda}$ también es autovalor de A. Es decir, los autovalores de A no reales son conjugados a pares.
   c) Si n es impar, A tiene por lo menos un autovalor real.
3. Sea V un espacio euclídeo real de dimensión n. Una transformación ortogonal $T: V \to V$ con determinante igual a 1 se llama *rotación*. Si n es impar, demostrar que 1 es autovalor para T. Esto prueba que toda rotación de un espacio de dimensión impar tiene un eje fijo. [*Indicación:* Aplicar el ejercicio 2.]
4. Dada una matriz real ortogonal A con -1 como autovalor de multiplicidad k. Demostrar que $\det A = (-1)^k$.
5. Si T es lineal y conserva la norma, demostrar que T es unitaria.
6. Si $T: V \to V$ es unitaria y hermitiana, demostrar que $T^2 = I$.
7. Sean $(e_1, \ldots, e_n)$ y $(u_1, \ldots, u_n)$ dos bases ortonormales para un espacio euclídeo V. Demostrar que existe una transformación unitaria T que aplica una de esas bases en la otra.
8. Hallar un número real a tal que la siguiente matriz sea unitaria:

$$\begin{bmatrix} a & \frac{1}{2}i & \frac{1}{2}a(2i-1) \\ ia & \frac{1}{2}(1+i) & \frac{1}{2}a(1-i) \\ a & -\frac{1}{2} & \frac{1}{2}a(2-i) \end{bmatrix}.$$

9. Si A es una matriz hemi-hermitiana, demostrar que $I - A$ e $I + A$ son no singulares y que $(I - A)(I + A)^{-1}$ es unitaria.
10. Si A es una matriz unitaria y si $I + A$ es no singular, demostrar que $(I - A)(I + A)^{-1}$ es hemi-hermitiana.
11. Si A es hermitiana, demostrar que $A - iI$ es no singular y que $(A - iI)^{-1}(A + iI)$ es unitaria.
12. Demostrar que cualquier matriz unitaria puede diagonalizarse con una matriz unitaria.
13. Una matriz cuadrada se llama *normal* si $AA^* = A^*A$. Determinar cuál de los tipos siguientes de matrices son normales.
   a) Matrices hermitianas.
   b) Matrices hemi-hermitianas.
   c) Matrices simétricas.
   d) Matrices hemi-simétricas.
   e) Matrices unitarias.
   f) Matrices ortogonales.
14. Si A es una matriz normal ($AA^* = A^*A$) y si U es una matriz unitaria, demostrar que U^*AU es normal.

6

ECUACIONES DIFERENCIALES LINEALES

6.1 Introducción histórica

La historia de las ecuaciones diferenciales comenzó en el siglo XVII cuando Newton, Leibniz y los Bernoulli resolvieron algunas ecuaciones diferenciales sencillas de primero y segundo orden que se presentaron en problemas de Geometría y Mecánica. Estos primeros descubrimientos, iniciados alrededor de 1690, hicieron creer que las soluciones de todas las ecuaciones diferenciales originadas en problemas geométricos y físicos podrían expresarse por medio de las funciones elementales del Cálculo. Por ello, gran parte de los primeros esfuerzos fueron orientados al desarrollo de técnicas ingeniosas para resolver ecuaciones diferenciales por medio de recursos sencillos, como son la adición, sustracción, multiplicación, división, composición e integración, aplicadas tan sólo un número finito de veces a las funciones ordinarias del Cálculo.

Los métodos especiales, tales como la separación de variables y el empleo de factores integrantes, fueron ideados de manera más o menos casual antes de fines del siglo XVII. Durante el siglo XVIII, fueron desarrollados procedimientos más sistemáticos, principalmente por Euler, Lagrange y Laplace. Pronto se vio que relativamente pocas ecuaciones diferenciales podían resolverse con recursos elementales. Poco a poco, los matemáticos fueron dándose cuenta que era vano empeño el intentar descubrir métodos para resolver todas las ecuaciones diferenciales. En lugar de ello, encontraron más provechoso averiguar si una ecuación diferencial dada, tenía o no solución y, cuando la tenía, intentar la deducción de propiedades de la solución a partir de la misma ecuación diferencial. Con ello, los matemáticos empezaron a considerar las ecuaciones diferenciales como fuentes de nuevas funciones.

A principios del siglo XIX se desarrolló una fase importante de esa teoría, siguiendo una tendencia paralela a la de conseguir un desarrollo más riguroso del Cálculo. En 1820, Cauchy obtuvo el primer «teorema de existencia» para las

ecuaciones diferenciales. Demostró que toda ecuación de primer orden de la forma

$$y' = f(x, y)$$

tiene solución, siempre que el segundo miembro, $f(x, y)$, satisfaga ciertas condiciones generales. Un ejemplo importante es la ecuación de Ricatti

$$y' = P(x)y^2 + Q(x)y + R(x),$$

en la que P, Q y R son funciones dadas. El trabajo de Cauchy implica la existencia de una solución de la ecuación de Ricatti en cualquier intervalo abierto $(-r, r)$ en torno al origen, con tal que P, Q y R admitan desarrollos en serie de potencias en $(-r, r)$. En 1841 José Liouville (1809-1882) demostró que en algunos casos esa solución no puede obtenerse con medios elementales.

La experiencia ha puesto de manifiesto que es difícil obtener resultados de tipo general relativos a las soluciones de las ecuaciones diferenciales, salvo para unos pocos tipos. Entre éstos cabe citar las llamadas ecuaciones diferenciales *lineales* que se presentan en gran número de problemas científicos. En el Volumen I se estudiaron algunos tipos sencillos, las ecuaciones lineales de primer orden y las de segundo orden con coeficientes constantes. La siguiente sección se dedica a una revisión de los principales resultados relativos a esas ecuaciones.

6.2 Revisión de los resultados relativos a las ecuaciones de primero y segundo orden

Una ecuación diferencial lineal de primer orden es de la forma

$$y' + P(x)y = Q(x), \tag{6.1}$$

donde P y Q son funciones dadas. En el Volumen I demostramos un teorema de existencia y unicidad para esa ecuación (teorema 8.3) que aquí volvemos a enunciar.

TEOREMA 6.1. *Supongamos que P y Q son continuas en un intervalo abierto J. Elijamos cualquier punto a en J y sea b un número real cualquiera. Entonces existe una y sólo una función $y = f(x)$ que satisface la ecuación diferencial* (6.1) *y la condición inicial $f(a) = b$. Esa función viene dada por la fórmula*

$$f(x) = be^{-A(x)} + e^{-A(x)} \int_a^x Q(t)e^{A(t)}\, dt, \tag{6.2}$$

donde $A(x) = \int_a^x P(t)\, dt$.

Las ecuaciones lineales de segundo orden son las de la forma

$$P_0(x)y'' + P_1(x)y' + P_2(x)y = R(x).$$

Si los coeficientes P_0, P_1, P_2 y el segundo miembro R son continuas en un cierto intervalo J, y si P_0 no es nulo en ningún punto de J, un teorema (que se discute en la sección 6.5) garantiza que siempre existen soluciones en el intervalo J. Sin embargo, no existe una fórmula general análoga a la (6.2) para expresar esas soluciones en función de P_0, P_1, P_2 y R. Así pues, en esta generalización relativamente sencilla de (6.1), la teoría no es ni mucho menos completa, excepto en casos particulares. Si los coeficientes son *constantes* y si R es cero, todas las soluciones pueden determinarse explícitamente por medio de polinomios y funciones exponenciales y trigonométricas según afirma el siguiente teorema que se demostró en el Volumen I (teorema 8.7).

TEOREMA 6.2. *Consideremos la ecuación diferencial*

$$y'' + ay' + by = 0, \tag{6.3}$$

donde a y b son constantes reales dadas. Sea $d = a^2 - 4b$. Toda solución de (6.3) *en el intervalo $(-\infty, +\infty)$ tiene la forma*

$$y = e^{-ax/2}[c_1u_1(x) + c_2u_2(x)], \tag{6.4}$$

donde c_1 y c_2 son constantes, y las funciones u_1 y u_2 se determinan de acuerdo con el signo algebraico de d como sigue:

a) *Si $d = 0$, $u_1(x) = 1$ y $u_2(x) = x$.*

b) *Si $d > 0$, $u_1(x) = e^{kx}$ y $u_2(x) = e^{-kx}$, donde $k = \frac{1}{2}\sqrt{d}$.*

c) *Si $d < 0$, $u_1(x) = \cos kx$ y $u_2(x) = \operatorname{sen} kx$, donde $k = \frac{1}{2}\sqrt{-d}$*

El número $d = a^2 - 4b$ es el *discriminante* de la forma cuadrática

$$r^2 + ar + b = 0. \tag{6.5}$$

Ésta es la llamada *ecuación característica* de la ecuación diferencial (6.3). Sus raíces son

$$r_1 = \frac{-a + \sqrt{d}}{2}, \qquad r_2 = \frac{-a - \sqrt{d}}{2}.$$

El signo de d determina la naturaleza de esas raíces. Si $d > 0$ ambas raíces son reales y la solución de (6.4) puede expresarse en la forma

$$y = c_1 e^{r_1 x} + c_2 e^{r_2 x}.$$

Si $d < 0$, las raíces r_1 y r_2 son números complejos conjugados. Cada una de las funciones exponenciales complejas $f_1(x) = e^{r_1 x}$ y $f_2(x) = e^{r_2 x}$ es una solución compleja de la ecuación diferencial (6.3). Obtenemos soluciones reales examinando las partes real e imaginarias de f_1 y f_2. Escribiendo $r_1 = -\frac{1}{2}a + ik$, $r_2 = -\frac{1}{2}a - ik$, siendo $k = \frac{1}{2}\sqrt{-d}$, tenemos

$$f_1(x) = e^{r_1 x} = e^{-ax/2} e^{ikx} = e^{-ax/2} \cos kx + ie^{-ax/2} \operatorname{sen} kx$$

y

$$f_2(x) = e^{r_2 x} = e^{-ax/2} e^{-ikx} = e^{-ax/2} \cos kx - ie^{-ax/2} \operatorname{sen} kx.$$

La solución general que aparece en la ecuación (6.4) es una combinación lineal de las partes real e imaginaria de $f_1(x)$ y $f_2(x)$.

6.3 Ejercicios

Estos ejercicios han sido seleccionados del capítulo 8 del volumen I y constituyen una revisión sobre las ecuaciones diferenciales de primero y segundo orden.

Ecuaciones lineales de primer orden. En los ejercicios 1, 2, 3, resolver el problema de valores iniciales en el intervalo que se cita.

1. $y' - 3y = e^{2x}$ en $(-\infty, +\infty)$, con $y = 0$ cuando $x = 0$.
2. $xy' - 2y = x^5$ en $(0, +\infty)$, con $y = 1$ cuando $x = 1$.
3. $y' + y \tan x = \operatorname{sen} 2x$ en $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, con $y = 2$ cuando $x = 0$.
4. Si una cepa de bacterias crece proporcionalmente a la cantidad de individuos presente y si la población se dobla en una hora, ¿cuánto crecerá al cabo de dos horas?
5. Una curva de ecuación cartesiana $y = f(x)$ pasa por el origen. Por un punto arbitrario de la curva se trazan rectas paralelas a los ejes coordenados que forman un rectángulo con ellos. La curva divide todo rectángulo de esos en dos regiones A y B, una de las cuales tiene un área igual a n veces la otra. Hallar la función f.
6. a) Sea u una solución no nula de la ecuación de segundo orden $y'' + P(x)y' + Q(x)y = 0$. Demostrar que la sustitución $y = uv$ convierte la ecuación

$$y'' + P(x)y' + Q(x)y = R(x)$$

en una ecuación lineal de primer orden en v'.
b) Obtener una solución no nula de la ecuación $y'' - 4y' + x^2(y' - 4y) = 0$ por simple inspección y emplear el método de la parte a) para encontrar una solución de

$$y'' - 4y' + x^2(y' - 4y) = 2xe^{-x^3/3}$$

tal que $y = 0$ e $y' = 4$ cuando $x = 0$.

Ecuaciones lineales de segundo orden con coeficientes constantes. En cada uno de los ejercicios del 7 al 10, hallar todas las soluciones en $(-\infty, +\infty)$.

7. $y'' - 4y = 0$.
8. $y'' + 4y = 0$.
9. $y'' - 2y' + 5y = 0$.
10. $y'' + 2y' + y = 0$.

11. Hallar todos los valores de la constante k tales que la ecuación diferencial $y'' + ky = 0$ admita una solución no trivial $y = f_k(x)$ para la que $f_k(0) = f_k(1) = 0$. Para cada uno de esos valores de k, determinar la correspondiente solución $y = f_k(x)$. Considerar los valores de k positivos y negativos.
12. Si (a, b) es un punto dado del plano y m es un número real dado, demostrar que la ecuación diferencial $y'' + k^2y = 0$ tiene exactamente una solución cuya gráfica pasa por el punto (a, b) y tiene en él pendiente m. Discutir separadamente el caso $k = 0$.
13. En cada caso, hallar una ecuación diferencial lineal de segundo orden que se satisfaga con u_1 y u_2.

a) $u_1(x) = e^x, \quad u_2(x) = e^{-x}$.
b) $u_1(x) = e^{2x}, \quad u_2(x) = xe^{2x}$.
c) $u_1(x) = e^{-x/2}\cos x, \quad u_2(x) = e^{-x/2}\operatorname{sen} x$.
d) $u_1(x) = \operatorname{sen}(2x + 1), \quad u_2(x) = \operatorname{sen}(2x + 2)$.
e) $u_1(x) = \cosh x, \quad u_2(x) = \operatorname{senh} x$.

14. Una partícula está sometida a movimiento armónico simple. Inicialmente está situada en el punto 1, su velocidad es 1 y su aceleración es -12. Calcular su posición y aceleración cuando la velocidad sea $\sqrt{8}$

6.4 Ecuaciones diferenciales lineales de orden n

Una ecuación diferencial lineal de orden n es de la forma

$$P_0(x)y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = R(x). \tag{6.6}$$

Las funciones $P_0, P_1, \ldots, P_n$ que multiplican las diversas derivadas de la función incógnita y se llaman *coeficientes* de la ecuación. En nuestra discusión general de la ecuación lineal supondremos que todos los coeficientes son funciones continuas en un cierto intervalo J. La palabra «intervalo» se referirá a un intervalo acotado o no acotado.

En la ecuación diferencial (6.6) el coeficiente P_0 desempeña un papel especial, puesto que determina el orden de la ecuación. Los puntos en los que $P_0(x) = 0$ se llaman *puntos* singulares de la ecuación. La presencia de puntos singulares introduce, algunas veces, complicaciones que requieren un estudio especial. Para evitarlas supongamos que la función P_0 nunca es cero en J. Entonces podemos

dividir ambos miembros de la ecuación (6.6) por P_0 y escribir la ecuación diferencial con primer coeficiente igual a 1. Por consiguiente, en nuestra discusión general supondremos que la ecuación diferencial tiene la forma

$$y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = R(x). \tag{6.7}$$

El estudio de las ecuaciones lineales puede simplificarse mediante el uso de la notación de operadores. Designemos con $\mathscr{C}(J)$ el espacio lineal constituido por todas las funciones de valores reales, continuas en un intervalo J, y con $\mathscr{C}^n(J)$ el subespacio de todas las funciones f cuyas n primeras derivadas $f', f'', \ldots, f^{(n)}$ existen y son continuas en J. Sean $P_1, \ldots, P_n$, n funciones dadas en $\mathscr{C}(J)$ y consideremos el operador $L: \mathscr{C}^n(J) \to \mathscr{C}(J)$ dado por

$$L(f) = f^{(n)} + P_1 f^{(n-1)} + \cdots + P_n f.$$

El operador L a veces se escribe así:

$$L = D^n + P_1 D^{n-1} + \cdots + P_n,$$

donde D^k representa el operador derivada k-ésima. Con esa notación la ecuación diferencial (6.7) puede escribirse como sigue

$$L(y) = R. \tag{6.8}$$

Una solución de esta ecuación es cualquier función y de $\mathscr{C}^n(J)$ que satisfaga (6.8) en el intervalo J.

Es fácil comprobar que $L(y_1 + y_2) = L(y_1) + L(y_2)$, y que $L(cy) = cL(y)$ cualquiera que sea la constante c. Por lo tanto L es un operador *lineal*. Por ello la ecuación $L(y) = R$ se denomina lineal. El operador L se llama *operador diferencial lineal de orden n.*

A cada ecuación lineal $L(y) = R$ podemos asociar la ecuación

$$L(y) = 0,$$

en donde el segundo miembro se ha sustituido por cero. Esta se llama *ecuación homogénea* correspondiente a $L(y) = R$. Cuando R no es idénticamente cero, la ecuación $L(y) = R$ es una *ecuación no homogénea.* Encontraremos que siempre podemos resolver la ecuación no homogénea si nos es posible resolver la correspondiente ecuación homogénea. Por consiguiente, comenzaremos estudiando las homogéneas.

El conjunto de soluciones de la ecuación homogénea es el núcleo $N(L)$ del operador L. Éste se llama también el *espacio solución* de la ecuación. El espacio

solución es un subespacio de $\mathscr{C}^n(J)$. Si bien $\mathscr{C}^n(J)$ es de dimensión infinita, resulta que el espacio solución $N(L)$ siempre es de dimensión finita. En efecto, demostraremos que

$$\dim N(L) = n, \tag{6.9}$$

siendo n el orden del operador L. La ecuación (6.9) es el llamado *teorema de dimensionalidad* para los operadores diferenciales lineales. El teorema de dimensionalidad se deducirá como consecuencia de un teorema de existencia y unicidad que a continuación comentamos.

6.5 Teorema de existencia y unicidad

TEOREMA 6.3. TEOREMA DE EXISTENCIA Y UNICIDAD PARA ECUACIONES LINEALES DE ORDEN n. *Sean $P_1, P_2, \ldots, P_n$ funciones continuas en un intervalo abierto J, y sea L el operador diferencial lineal*

$$L = D^n + P_1 D^{n-1} + \cdots + P_n .$$

Si $x_0 \in J$ y si $k_0, k_1, \ldots, k_{n-1}$ son n números reales dados, existe entonces una y sólo una función $y = f(x)$ que satisface la ecuación diferencial homogénea $L(y) = 0$ en J y que también satisface las condiciones iniciales

$$f(x_0) = k_0, f'(x_0) = k_1, \ldots, f^{(n-1)}(x_0) = k_{n-1} .$$

Observación: El vector del n-espacio dado por $(f(x_0), f'(x_0), \ldots, f^{(n-1)}(x_0))$ se llama *vector-valor inicial* de f en x_0. El teorema 6.3 nos dice que si elegimos un punto x_0 en J y un vector en el n-espacio, la ecuación homogénea $L(y) = 0$ tiene exactamente una solución $y = f(x)$ en J con aquel vector como valor inicial en x_0. Por ejemplo, cuando $n = 2$ existe exactamente una solución con valor asignado $f(x_0)$ y derivada asignada $f'(x_0)$ en el punto x_0.

La demostración del teorema de existencia y unicidad se obtendrá como corolario de los teoremas de existencia y unicidad más generales que se estudian en el capítulo 7. En la sección 7.9 se estudia otra demostración para el caso de ecuaciones con coeficientes constantes.

6.6 Dimensión del espacio solución de una ecuación lineal homogénea

TEOREMA 6.4. TEOREMA DE DIMENSIONALIDAD. *Sea $L: \mathscr{C}^n(J) \to \mathscr{C}(J)$ un operador diferencial lineal de orden n dado por*

$$L = D^n + P_1 D^{n-1} + \cdots + P_n . \tag{6.10}$$

Entonces el espacio solución de la ecuación $L(y) = 0$ tiene dimensión n.

Demostración. Sea V_n el espacio lineal n-dimensional de las n-plas de escalares. Sea T la transformación lineal que aplica cada función f del espacio solución $N(L)$ en el vector-valor inicial de f en x_0,

$$T(f) = (f(x_0), f'(x_0), \ldots, f^{(n-1)}(x_0)),$$

siendo x_0 un punto fijo en J. El teorema de unicidad nos dice que $T(f) = 0$ implica que $f = 0$. Por consiguiente, según el teorema 2.10, T es uno a uno en $N(L)$. Luego T^{-1} también es uno a uno y aplica V_n en $N(L)$, y el teorema 2.11 demuestra que $\dim N(L) = \dim V_n = n$.

Puesto que el espacio solución tiene dimensión n, un conjunto cualquiera de n soluciones independientes servirá de base. Por consiguiente, como corolario del teorema de dimensionalidad tenemos:

TEOREMA 6.5. *Sea $L: \mathscr{C}^n(J) \rightarrow \mathscr{C}(J)$ un operador diferencial lineal de orden n. Si $u_1, \ldots, u_n$ son n soluciones independientes de la ecuación diferencial homogénea $L(y) = 0$ en J, toda solución $y = f(x)$ en J puede expresarse en la forma*

(6.11) $$f(x) = \sum_{k=1}^{n} c_k u_k(x),$$

siendo $c_1, \ldots, c_n$ constantes.

Observación: Puesto que todas las soluciones de la ecuación diferencial $L(y) = 0$ están contenidas en la fórmula (6.11), la combinación lineal del segundo miembro, con las constantes arbitrarias $c_1, \ldots, c_n$, a veces se llama *solución general* de la ecuación diferencial.

El teorema de dimensionalidad nos dice que el espacio solución de una ecuación diferencial lineal homogénea de orden n tiene siempre una base de n soluciones, pero no nos dice cómo se determina esa base. En realidad, no se conoce un método sencillo para determinar una base de soluciones para toda ecuación lineal. No obstante, se han ideado métodos especiales para casos particulares. Entre éstos cabe considerar las ecuaciones diferenciales con coeficientes constantes que volvemos ahora a considerar.

6.7 Álgebra de operadores de coeficientes constantes

Un operador A de coeficientes constantes es un operador lineal de la forma

(6.12) $$A = a_0 D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n,$$

donde D es el operador derivada y $a_0, a_1, \ldots, a_n$ son constantes reales. Si $a_0 \neq 0$ el operador se llama de orden n. El operador A puede aplicarse a cualquier función y con n derivadas en un cierto intervalo, siendo el resultado una función $A(y)$ dada por

$$A(y) = a_0 y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y .$$

En esta sección, limitamos nuestro estudio a funciones que tienen derivadas de cualquier orden en $(-\infty, +\infty)$. El conjunto de tales funciones se designará con $\mathscr{C}^\infty$ y se le llamará clase de las *funciones indefinidamente derivables.* Si $y \in \mathscr{C}^\infty$, entonces $A(y)$ pertenece también a $\mathscr{C}^\infty$.

Las operaciones algebraicas corrientes con las transformaciones lineales (*adición, multiplicación por escalares,* y *composición* o *multiplicación*) pueden aplicarse, en particular, a los operadores con coeficientes constantes. Sean A y B dos operadores con coeficientes constantes (no necesariamente del mismo orden). Puesto que la suma $A + B$ y todos los productos λA son también operadores con coeficientes constantes, el conjunto de todos los operadores con coeficientes constantes es un espacio lineal. El producto de A por B (en cualquier orden) es también un operador con coeficientes constantes. Por consiguiente, las sumas, productos y productos por escalares de operadores con coeficientes constantes satisfacen las leyes conmutativa, asociativa y distributiva ordinarias que cumplen todas las transformaciones lineales. Asimismo, puesto que $D^r D^s = D^s D^r$ para todos los enteros positivos r y s, dos operadores con coeficientes constantes cualesquiera A y B son permutables; $AB = BA$.

A cada operador con coeficientes constantes A asociamos un polinomio p_A llamado polinomio característico de A. Si A viene dado por (6.12), p_A es el polinomio que tiene los mismos coeficientes que A. Esto es, para todo r real tenemos

$$p_A(r) = a_0 r^n + a_1 r^{n-1} + \cdots + a_n .$$

Recíprocamente, dado un polinomio real cualquiera p, existe un correspondiente operador A cuyos coeficientes son los mismos que los de p. El siguiente teorema muestra que tal asociación entre operadores y polinomios es una correspondencia uno a uno. Además, esta correspondencia asocia a sumas, productos de operadores y productos de escalares por operador las respectivas sumas, productos y productos por escalares de sus polinomios característicos.

TEOREMA 6.6. *Sean A y B dos operadores con coeficientes constantes con polinomios característicos respectivos p_A y p_B, y λ un número real. Entonces tenemos:*

a) $A = B$ *si y sólo si* $p_A = p_B$,

b) $p_{A+B} = p_A + p_B$,

c) $p_{AB} = p_A \cdot p_B$,

d) $p_{\lambda A} = \lambda \cdot p_A$.

Demostración. Consideremos primero la parte a). Supongamos $p_A = p_B$. Queremos demostrar que $A(y) = B(y)$ para toda y de $\mathscr{C}^\infty$. Puesto que $p_A = p_B$, ambos polinomios tienen el mismo grado y los mismos coeficientes. Por consiguiente A y B tienen el mismo orden y los mismos coeficientes, así que $A(y) = B(y)$ para toda y de $\mathscr{C}^\infty$.

Seguidamente demostramos que $A = B$ implica $p_A = p_B$. La relación $A = B$ significa que $A(y) = B(y)$ para toda y de $\mathscr{C}^\infty$. Tomemos $y = e^{rx}$, siendo r una constante. Puesto que $y^{(k)} = r^k e^{rx}$ para todo $k \geq 0$, tenemos

$$A(y) = p_A(r)e^{rx} \quad \text{y} \quad B(y) = p_B(r)e^{rx}.$$

La ecuación $A(y) = B(y)$ implica $p_A(r) = p_B(r)$. Ya que r es cualquiera, debe ser $p_A = p_B$. Esto completa la demostración de la parte a).

Las partes b), c) y d) se deducen al momento de la definición de polinomio característico.

Del teorema 6.6 resulta que toda relación algebraica que incluya sumas, productos y productos por escalares de los polinomios p_A y p_B son también válidos para los operadores A y B. En particular, si el polinomio característico p_A puede factorizarse como producto de dos o más polinomios, cada factor debe ser el polinomio característico de un cierto operador con coeficientes constantes, así que según el teorema 6.6 existe una correspondiente factorización del operador A. Por ejemplo, si $p_A(r) = p_B(r)p_C(r)$, entonces $A = BC$. Si $p_A(r)$ puede factorizarse como producto de n factores lineales, por ejemplo sea

$$p_A(r) = a_0(r - r_1)(r - r_2) \cdots (r - r_n). \tag{6.13}$$

la correspondiente factorización de A toma la forma

$$A = a_0(D - r_1)(D - r_2) \cdots (D - r_n).$$

El teorema fundamental del álgebra nos dice que todo polinomio $p_A(r)$ de grado $n \geq 1$ tiene una factorización de la forma (6.13), donde $r_1, r_2, \ldots, r_n$ son las raíces de la ecuación

$$p_A(r) = 0,$$

llamada *ecuación característica* de A. Cada raíz se escribe tantas veces como indica su orden de multiplicidad. Las raíces pueden ser reales o complejas. Puesto que $p_A(r)$ tiene coeficientes reales, las raíces complejas se presentan en pares de conjugadas, $\alpha + i\beta$, $\alpha - i\beta$, si $\beta \neq 0$. Los dos factores lineales correspondientes a cada uno de tales pares puede combinarse para dar un factor cuadrático

$r^2 - 2\alpha r + \alpha^2 + \beta^2$ cuyos coeficientes son reales. Por consiguiente, todo polinomio $p_A(r)$ puede factorizarse como un producto de polinomios lineales y cuadráticos *con coeficientes reales.* Esto nos proporciona una correspondiente factorización del operador A como un producto de operadores de primer y segundo orden con coeficientes constantes y reales.

EJEMPLO 1. Sea $A = D^2 - 5D + 6$. Puesto que el polinomio característico $p_A(r)$ tiene la factorización $r^2 - 5r + 6 = (r - 2)(r - 3)$, el operador A tiene la factorización

$$D^2 - 5D + 6 = (D - 2)(D - 3).$$

EJEMPLO 2. Sea $A = D^4 - 2D^3 + 2D^2 - 2D + 1$. El polinomio característico $p_A(r)$ tiene la factorización

$$r^4 - 2r^3 + 2r^2 - 2r + 1 = (r - 1)(r - 1)(r^2 + 1),$$

así que A tiene la factorización $A = (D - 1)(D - 1)(D^2 + 1)$.

6.8 Determinación de una base de soluciones para ecuaciones lineales con coeficientes constantes por factorización de operadores

El teorema que sigue demuestra como la factorización de operadores con coeficientes constantes nos ayuda en la resolución de ecuaciones diferenciales lineales con coeficientes constantes.

TEOREMA 6.7. *Sea L un operador con coeficientes constantes que puede factorizarse como un producto de operadores con coeficientes constantes, por ejemplo*

$$L = A_1 A_2 \cdots A_k .$$

Entonces el espacio solución de la ecuación diferencial lineal $L(y) = 0$ contiene el espacio solución de cada una de las ecuaciones diferenciales $A_i(y) = 0$. Dicho de otro modo,

$$N(A_i) \subseteq N(L) \quad \textit{para cada } i = 1, 2, \ldots, k. \tag{6.14}$$

Demostración. Si u es el núcleo del último factor A_k tenemos $A_k(u) = 0$, así que

$$L(u) = (A_1 A_2 \cdots A_k)(u) = (A_1 \cdots A_{k-1}) A_k(u) = (A_1 \cdots A_{k-1})(0) = 0.$$

Por lo tanto, el núcleo de L contiene el núcleo del último factor A_k. Pero ya que los operadores con coeficientes constantes son permutables, podemos reordenar los factores de modo que uno cualquiera de ellos sea el último factor. Esto demuestra (6.14).

Si $L(u) = 0$, se dice que el operador L *anula* u. El teorema 6.7 nos dice que si un factor A_i de L anula u, también L anula u.

Vamos a ver cómo puede utilizarse el teorema para resolver las ecuaciones diferenciales homogéneas con coeficientes constantes. Hemos elegido ejemplos para ilustrar distintos aspectos, que dependen de la naturaleza de las raíces de la ecuación característica.

CASO I. Raíces reales distintas.

EJEMPLO 1. Hallar una base de soluciones para la ecuación diferencial

$$(D^3 - 7D + 6)y = 0. \tag{6.15}$$

Solución. Tiene la forma $L(y) = 0$ con

$$L = D^3 - 7D + 6 = (D - 1)(D - 2)(D + 3).$$

El núcleo de $D - 1$ contiene $u_1(x) = e^x$, el de $D - 2$ contiene $u_2(x) = e^{2x}$, y el de $D + 3$ contiene $u_3(x) = e^{-3x}$. En el capítulo 1 (p. 12) se demostró que u_1, u_2, u_3 son independientes. Puesto que tres soluciones independientes de una ecuación de tercer orden forman una base para el espacio solución, la solución general de (6.15) viene dada por

$$y = c_1e^x + c_2e^{2x} + c_3e^{-3x}.$$

El método utilizado para resolver el ejemplo 1 nos permite encontrar una base para el espacio solución de cualquier operador con coeficientes constantes que pueda descomponerse en producto de factores lineales distintos.

TEOREMA 6.8. *Sea L un operador con coeficientes constantes cuya ecuación característica $p_L(r) = 0$ tiene n raíces reales distintas $r_1, r_2, \ldots, r_n$. Entonces la solución general de la ecuación diferencial $L(y) = 0$ en el intervalo $(-\infty, +\infty)$ viene dada por la fórmula*

$$y = \sum_{k=1}^{n} c_k e^{r_k x}. \tag{6.16}$$

Demostración. Tenemos la factorización

$$L = a_0(D - r_1)(D - r_2) \cdots (D - r_n).$$

Puesto que el núcleo de $(D - r_k)$ contiene $u_k(x) = e^{r_k x}$, el núcleo de L contiene las n funciones

$$u_1(x) = e^{r_1 x}, \qquad u_2(x) = e^{r_2 x}, \ldots, u_n(x) = e^{r_n x}.$$

En el capítulo 1 (p. 12) se demostró que esas funciones son independientes. Por consiguiente constituyen una base para el espacio solución de la ecuación $L(y)=0$, así que la solución general viene dada por (6.16).

CASO II. Raíces reales múltiples.

Si todas las raíces son reales pero no distintas, las funciones (6.17) no son independientes y por tanto no forman una base para el espacio solución. Si una raíz r se presenta con multiplicidad m, entonces $(D - r)^m$ es un factor de L. El teorema que sigue nos dice cómo obtener m soluciones independientes en el núcleo de ese factor.

TEOREMA 6.9. *Las m funciones*

$$u_1(x) = e^{rx}, \qquad u_2(x) = xe^{rx}, \ldots, u_m(x) = x^{m-1}e^{rx} \tag{6.17}$$

son m elementos independientes anulados por el operador $(D - r)^m$.

Demostración. La independencia de esas funciones se deduce de la independencia de los polinomios $1, x, x^2, \ldots, x^{m-1}$. Para demostrar que $u_1, u_2, \ldots, u_m$ son anuladas por $(D - r)^m$ lo haremos por inducción.

Si $m = 1$ existe tan sólo una función, $u_1(x) = e^{rx}$, que evidentemente es anulada por $(D - r)$. Supongamos entonces, que el teorema es cierto para $m - 1$. Esto significa que las funciones $u_1, \ldots, u_{m-1}$ son anuladas por $(D - r)^{m-1}$. Puesto que

$$(D - r)^m = (D - r)(D - r)^{m-1}$$

las funciones $u_1, \ldots, u_{m-1}$ también son anuladas por $(D - r)^m$. Para completar la demostración tenemos que demostrar que $(D - r)^m$ anula u_m. Consideremos por tanto

Tenemos

$$(D - r)^m u_m = (D - r)^{m-1}(D - r)(x^{m-1}e^{rx}).$$

$$\begin{aligned}(D - r)(x^{m-1}e^{rx}) &= D(x^{m-1}e^{rx}) - rx^{m-1}e^{rx} \\ &= (m - 1)x^{m-2}e^{rx} + x^{m-1}re^{rx} - rx^{m-1}e^{rx} \\ &= (m - 1)x^{m-2}e^{rx} = (m - 1)u_{m-1}(x).\end{aligned}$$

Cuando aplicamos $(D - r)^{m-1}$ a los dos miembros de esta última ecuación obtenemos 0 en el segundo miembro ya que $(D - r)^{m-1}$ anula a u_{m-1}. Luego $(D - r)^m u_m = 0$ con lo cual u_m es anulada por $(D - r)^m$. Esto completa la demostración.

EJEMPLO 2. Hallar la solución general de la ecuación diferencial $L(y) = 0$, donde $L = D^3 - D^2 - 8D + 12$.

Solución. El operador L tiene la factorización

$$L = (D - 2)^2(D + 3).$$

Según el teorema 6.9, las dos funciones

$$u_1(x) = e^{2x}, \qquad u_2(x) = xe^{2x}$$

pertenecen al núcleo de $(D - 2)^2$. La función $u_3(x) = e^{-3x}$ pertenece al núcleo de $(D + 3)$. Como quiera que u_1, u_2, u_3 son independientes (ver ejercicio 17 de la sección 6.9) forman una base para el núcleo de L, así que la solución general de la ecuación diferencial es

$$y = c_1e^{2x} + c_2xe^{2x} + c_3e^{-3x}.$$

El teorema 6.9 nos indica cómo hallar una base de soluciones para cualquier ecuación lineal de orden n con coeficientes constantes cuya ecuación característica tenga sólo raíces reales, alguna de las cuales se repita. Si las raíces *distintas* son $r_1, r_2, \ldots, r_k$ y se presentan con las multiplicidades respectivas $m_1, m_2, \ldots, m_k$, la parte de la base que corresponde a r_p viene dada por las m_p funciones

$$u_{q,p}(x) = x^{q-1}e^{r_px}, \qquad \text{donde} \quad q = 1, 2, \ldots, m_p.$$

Cuando p toma los valores $1, 2, \ldots, k$ obtenemos en total $m_1 + \ldots + m_k$ funciones. En el ejercicio 17 de la sección 6.9 esbozamos una demostración haciendo ver que todas esas funciones son independientes. Puesto que la suma de las mul-

tiplicidades $m_1 + \ldots + m_k$ es igual a n, el orden de la ecuación, las funciones $u_{p,q}$ forman una base para el espacio solución de la ecuación.

EJEMPLO 3. Resolver la ecuación $(D^6 + 2D^5 - 2D^3 - D^2)y = 0$.

Solución. Tenemos $D^6+2D^5-2D^3-D^2=D^2(D-1)(D+1)^3$. La parte de la base correspondiente al factor D^2 es $u_1(x) = 1$, $u_2(x) = x$; la parte correspondiente al factor $(D - 1)$ es $u_3(x) = e^x$; y la correspondiente al factor $(D + 1)^3$ es $u_4(x) = e^{-x}$, $u_5(x) = xe^{-x}$, $u_6(x) = x^2e^{-x}$. Las seis funciones $u_1, \ldots, u_6$ son independientes, así que la solución general de la ecuación es

$$y = c_1 + c_2x + c_3e^x + (c_4 + c_5x + c_6x^2)e^{-x}.$$

CASO III. Raíces complejas.

Si se emplean exponenciales complejas, no hay necesidad de distinguir entre raíces reales y complejas de la ecuación característica de la ecuación diferencial $L(y) = 0$. Si se desean soluciones de valor real, factorizamos el operador L en factores lineales y cuadráticos con coeficientes reales. Cada par de raíces complejas conjugadas $\alpha + i\beta$, $\alpha - i\beta$ corresponde al factor cuadrático,

$$D^2 - 2\alpha D + \alpha^2 + \beta^2. \tag{6.18}$$

El núcleo de este operador de segundo orden contiene las dos funciones independientes $u(x) = e^{\alpha x} \cos \beta x$ y $v(x) = e^{\alpha x} \operatorname{sen} \beta x$. Si el par de raíces $\alpha \pm i\beta$ se presenta con multiplicidad m, el factor cuadrático se presenta elevado a la potencia m-ésima. El núcleo del operador

$$[D^2 - 2\alpha D + \alpha^2 + \beta^2]^m$$

contiene $2m$ funciones independientes,

$$u_q(x) = x^{q-1}e^{\alpha x} \cos \beta x, \qquad v_q(x) = x^{q-1}e^{\alpha x} \operatorname{sen} \beta x, \qquad q = 1, 2, \ldots, m.$$

Esto puede probarse fácilmente por inducción respecto a m. (En el ejercicio 20 de la sección 6.9 se da un esquema de las demostraciones.) Los ejemplos siguientes ilustran algunas de las posibilidades.

EJEMPLO 4. $y''' - 4y'' + 13y' = 0$. La ecuación característica, $r^3 - 4r^2 + 13r = 0$, tiene las raíces $0, 2 \pm 3i$; la solución general es

$$y = c_1 + e^{2x}(c_2 \cos 3x + c_3 \operatorname{sen} 3x).$$

EJEMPLO 5. $y''' - 2y'' + 4y' - 8y = 0$. La ecuación característica es

$$r^3 - 2r^2 + 4r - 8 = (r-2)(r^2+4) = 0\,;$$

sus raíces son $2, 2i, -2i$, así que la solución general de la ecuación diferencial es

$$y = c_1e^{2x} + c_2\cos 2x + c_3\operatorname{sen} 2x\,.$$

EJEMPLO 6. $y^{(5)} - 9y^{(4)} + 34y''' - 66y'' + 65y' - 25y = 0$. La ecuación característica puede escribirse así

$$(r-1)(r^2-4r+5)^2 = 0\,;$$

sus raíces son $1, 2 \pm i, 2 \pm i$, por lo que la solución general de la ecuación diferencial es

$$y = c_1e^{x} + e^{2x}[(c_2 + c_3x)\cos x + (c_4 + c_5x)\operatorname{sen} x]\,.$$

6.9 Ejercicios

Hallar la solución general de cada una de las ecuaciones diferenciales de los ejercicios 1 al 12.

1. $y''' - 2y'' - 3y' = 0$.
2. $y''' - y' = 0$.
3. $y''' + 4y'' + 4y' = 0$.
4. $y''' - 3y'' + 3y' - y = 0$.
5. $y^{(4)} + 4y''' + 6y'' + 4y' + y = 0$.
6. $y^{(4)} - 16y = 0$.
7. $y^{(4)} + 16y = 0$.
8. $y''' - y = 0$.
9. $y^{(4)} + 4y''' + 8y'' + 8y' + 4y = 0$.
10. $y^{(4)} + 2y'' + y = 0$.
11. $y^{(6)} + 4y^{(4)} + 4y'' = 0$.
12. $y^{(6)} + 8y^{(4)} + 16y'' = 0$.

13. Si m es una constante positiva, hallar la solución particular $y = f(x)$ de la ecuación diferencial

$$y''' - my'' + m^2y' - m^3y = 0$$

que satisfaga las condiciones $f(0) = f'(0) = 0$, $f''(0) = 1$.

14. Una ecuación diferencial lineal con coeficientes constantes tiene la ecuación característica $f(r) = 0$. Si todas las raíces de la ecuación característica son negativas, demostrar que toda solución de la ecuación diferencial tiende a cero cuando $x \to +\infty$. ¿Qué se puede asegurar acerca del comportamiento de todas las soluciones en el intervalo $[0, +\infty)$ si todas las raíces de la ecuación característica son no positivas?

15. En cada caso, hallar una ecuación diferencial lineal con coeficientes constantes que se satisfaga para todas las funciones dadas,

a) $u_1(x) = e^x$, $u_2(x) = e^{-x}$, $u_3(x) = e^{2x}$, $u_4(x) = e^{-2x}$.
b) $u_1(x) = e^{-2x}$, $u_2(x) = xe^{-2x}$, $u_3(x) = x^2e^{-2x}$.
c) $u_1(x) = 1$, $u_2(x) = x$, $u_3(x) = e^x$, $u_4(x) = xe^x$.
d) $u_1(x) = x$, $u_2(x) = e^x$, $u_3(x) = xe^x$.
e) $u_1(x) = x^2$, $u_2(x) = e^x$, $u_3(x) = xe^x$.
f) $u_1(x) = e^{-2x}\cos 3x$, $u_2(x) = e^{-2x}\operatorname{sen} 3x$, $u_3(x) = e^{-2x}$, $u_4(x) = xe^{-2x}$.
g) $u_1(x) = \cosh x$, $u_2(x) = \operatorname{senh} x$, $u_3(x) = x\cosh x$, $u_4(x) = x\operatorname{senh} x$.
h) $u_1(x) = \cosh x \operatorname{sen} x$, $u_2(x) = \operatorname{senh} x \cos x$, $u_3(x) = x$.

16. Sean $r_1, \ldots, r_n$, n números reales distintos, y $Q_1, \ldots, Q_n$, n polinomios, ninguno de los cuales es el polinomio nulo. Demostrar que las n funciones

$$u_1(x) = Q_1(x)e^{r_1x}, \ldots, u_n(x) = Q_n(x)e^{r_nx}$$

son independientes.

Bosquejo de la demostración. Emplear la inducción respecto a n. Para $n = 1$ y $n = 2$ el resultado es fácilmente comprobable. Supongamos que la proposición es cierta para $n = p$ y sean $c_1, \ldots, c_{p+1}$, $p + 1$ escalares reales tales que

$$\sum_{k=1}^{p+1} c_kQ_k(x)e^{r_kx} = 0.$$

Multipliquemos los dos miembros por $e^{-r_{p+1}x}$ y derivemos la ecuación resultante. Emplear entonces la hipótesis de inducción para probar que todos los escalares c_k son 0. Puede darse otra demostración basada en el orden de magnitud cuando $x \to +\infty$, como se hizo en el ejemplo 7 de la sección 1.7 (p. 12).

17. Sean $m_1, m_2, \ldots, m_k$, k enteros positivos, y $r_1, r_2, \ldots, r_k$, k números reales distintos, y sea $n = m_1 + \ldots + m_k$. Para cada par de enteros p, q que cumplan $1 \le p \le k$, $1 \le q \le m_p$, pongamos

$$u_{q,p}(x) = x^{q-1}e^{r_px}.$$

Por ejemplo, cuando $p = 1$ las correspondientes funciones son

$$u_{1,1}(x) = e^{r_1x}, \qquad u_{2,1}(x) = xe^{r_1x}, \ldots, u_{m_1,1}(x) = x^{m_1-1}e^{r_1x}.$$

Demostrar que las n funciones $u_{q,p}$ así definidas son independientes.

[*Indicación:* Aplicar el ejercicio 16.]

18. Sea L un operador diferencial lineal de orden n con coeficientes constantes y polinomio característico $p(r)$. Sea L' el operador de coeficientes constantes cuyo polinomio característico es la derivada $p'(r)$. Por ejemplo, si $L = 2D^2 - 3D + 1$ entonces $L' = 4D - 3$. Con mayor generalidad, definamos la derivada m-ésima $L^{(m)}$ como el operador cuyo polinomio característico es la derivada m-ésima $p^{(m)}(r)$. (El operador $L^{(m)}$ no debe confundirse con la potencia m-ésima L^m.)

a) Si u tiene n derivadas, demostrar que

$$L(u) = \sum_{k=0}^{n} \frac{p^{(k)}(0)}{k!} u^{(k)} .$$

b) Si u tiene $n - m$ derivadas, demostrar que

$$L^{(m)}(u) = \sum_{k=0}^{n-m} \frac{p^{(k+m)}(0)}{k!} u^{(k)} \qquad \text{para} \qquad m = 0, 1, 2, \ldots, n,$$

donde $L^{(0)} = L$.

19. Con la notación del ejercicio 18. Si u y v tienen n derivadas, demostrar que

$$L(uv) = \sum_{k=0}^{n} \frac{L^{(k)}(u)}{k!} v^{(k)} .$$

[*Indicación:* Emplear el ejercicio 18 junto con la fórmula de Leibniz para la derivada k-ésima de un producto:

$$(uv)^{(k)} = \sum_{r=0}^{k} \binom{k}{r} u^{(k-r)} v^{(r)} .]$$

20. a) Sea $p(t) = q(t)^m r(t)$, donde q y r son polinomios y m es un entero positivo. Demostrar que $p'(t) = q(t)^{m-1} s(t)$, donde s es un polinomio.
b) Sea L un operador con coeficientes constantes que anula u, siendo u una función dada de x. Sea $M = L^m$, la potencia m-ésima de L, donde $m > 1$. Demostrar que cada una de las derivadas $M', M'', \ldots, M^{(m-1)}$ anulan también u.
c) Con la parte b) y el ejercicio 19 demostrar que M anula cada una de las funciones $u, xu, \ldots, x^{m-1}u$.
d) Con la parte c) demostrar que el operador $(D^2 - 2\alpha D + \alpha^2 + \beta^2)^m$ anula cada una de las funciones $x^q e^{\alpha x} \operatorname{sen} \beta x$ y $x^q e^{\alpha x} \cos \beta x$ para $q = 1, 2, \ldots, m - 1$.

21. Sea L un operador de orden n con coeficientes constantes y polinomio característico $p(r)$. Si α es constante y u tiene n derivadas, demostrar que

$$L(e^{\alpha x} u(x)) = e^{\alpha x} \sum_{k=0}^{n} \frac{p^{(k)}(\alpha)}{k!} u^{(k)}(x) .$$

6.10 Relación entre las ecuaciones homogéneas y no homogéneas

Volvamos ahora a la ecuación diferencial lineal general de orden n con coeficientes que no sean necesariamente constantes. El teorema que sigue describe la relación entre las soluciones de una ecuación homogénea $L(y) = 0$ y las de una no homogénea $L(y) = R(x)$.

TEOREMA 6.10. *Sea $L : \mathscr{C}_n(J) \to \mathscr{C}(J)$ un operador diferencial lineal de orden*

n. Sean $u_1, \ldots, u_n$, n soluciones independientes de la ecuación homogénea $L(y)=0$, y sea y_1 una solución particular de la ecuación no homogénea $L(y) = R$, donde $R \in \mathscr{C}(J)$. Entonces toda solución $y = f(x)$ de la ecuación no homogénea tiene la forma

$$f(x) = y_1(x) + \sum_{k=1}^{n} c_k u_k(x), \tag{6.19}$$

donde $c_1, \ldots, c_n$ son constantes.

Demostración. Por la linealidad tenemos $L(f-y_1)=L(f)-L(y_1)=R-R=0$. Por lo tanto $f - y_1$ pertenece al espacio solución de la ecuación homogénea $L(y) = 0$, así que $f - y_1$ es una combinación lineal de $u_1, \ldots, u_n$, es decir, $f - y_1 = c_1u_1 + \ldots + c_nu_n$. Esto demuestra (6.19).

Puesto que todas las soluciones de $L(y) = R$ se obtienen con (6.19), la suma del segundo miembro de (6.19) (con las constantes arbitrarias $c_1, c_2, \ldots, c_n$) se llama la solución general de la ecuación no homogénea. El teorema 6.10 establece que la solución general de la ecuación no homogénea se obtiene sumando a y_1 la solución general de la ecuación homogénea.

Observación: El teorema 6.10 tiene una sencilla analogía geométrica que nos ayuda a profundizar en su significado. Para determinar todos los puntos de un plano encontramos un punto del mismo y le sumamos todos los puntos de un plano paralelo que pasa por el origen. Para encontrar todas las soluciones de $L(y) = R$ encontramos una solución particular y le sumamos todas las soluciones de la ecuación homogénea $L(y) = 0$. El conjunto de soluciones de la ecuación no homogénea es análogo al plano que pasa por un determinado punto. El espacio solución de la ecuación homogénea es análogo al plano paralelo que pasa por el origen.

En la práctica para utilizar el teorema 6.10 tenemos que resolver dos problemas: 1) Hallar la solución general de la ecuación homogénea $L(y) = 0$, y 2) hallar una solución particular de la ecuación no homogénea $L(y) = R$. En la próxima sección demostramos que siempre se puede resolver el problema 2) si podemos resolver el problema 1).

6.11 Determinación de una solución particular de la ecuación no homogénea. Método de variación de constantes

Volvamos a considerar el problema de la determinación de una solución particular y_1 de la ecuación no homogénea $L(y) = R$. Expondremos un método llamado de *variación de constantes* que nos indica cómo determinar y_1 si conocemos n soluciones independientes $u_1, \ldots, u_n$ de la ecuación homogénea $L(y) = 0$.

El método proporciona una solución particular de la forma

(6.20) $$y_1 = v_1u_1 + \cdots + v_nu_n,$$

donde $v_1, \ldots, v_n$ son funciones que pueden calcularse en función de $u_1, \ldots, u_n$ y del segundo miembro R. El método nos conduce a un sistema de n ecuaciones algebraicas lineales que se satisfacen con las derivadas $v_1', \ldots, v_n'$. Este sistema siempre puede resolverse debido a que tiene una matriz de coeficientes no singular. La integración de las derivadas, da entonces las funciones deseadas $v_1, \ldots, v_n$. El primero que utilizó el método fue Juan Bernoulli para resolver ecuaciones lineales de primer orden, y luego Lagrange en 1774 para resolver ecuaciones lineales de segundo orden.

Para el caso de orden n pueden simplificarse los detalles utilizando notación vectorial y matricial. El segundo miembro de (6.20) puede escribirse como un producto interior

(6.21) $$y_1 = (v, u),$$

donde v y u son funciones vectoriales n-dimensionales dadas por

$$v = (v_1, \ldots, v_n), \qquad u = (u_1, \ldots, u_n).$$

Intentamos elegir v de manera que el producto interior que define y_1 satisfaga la ecuación no homogénea $L(y) = R$, sabiendo que $L(u) = 0$, donde $L(u) = (L(u_1), \ldots, L(u_n))$.

Comenzamos calculando la primera derivada de y_1. Encontramos

(6.22) $$y_1' = (v, u') + (v', u).$$

Como tenemos n funciones $v_1, \ldots, v_n$ por determinar, deberemos precisar n condiciones que las relacionen. Si imponemos la condición de que el segundo término del segundo miembro de (6.22) se deba anular, la fórmula para y_1' se simplifica y queda

$$y_1' = (v, u'), \qquad \text{con tal que } (v', u) = 0.$$

Derivando la relación que da y_1' encontramos

$$y_1'' = (v, u'') + (v', u').$$

Si podemos elegir v de modo que $(v', u') = 0$ entonces la fórmula que da y_1'' también se simplifica y queda

$$y_1'' = (v, u''), \text{con tal que también} (v', u') = 0.$$

Si seguimos de este modo con las $n - 1$ primeras derivadas de y_1 encontramos

$$y_1^{(n-1)} = (v, u^{(n-1)}), \qquad \text{con tal que} \qquad (v', u^{(n-2)}) = 0.$$

Hasta aquí hemos establecido $n - 1$ condiciones a v. Derivando una vez más llegamos a

$$y_1^{(n)} = (v, u^{(n)}) + (v', u^{(n-1)}).$$

Impongamos ahora la condición $(v', u^{(n-1)}) = R(x)$, y la última ecuación se transforma en

$$y_1^{(n)} = (v, u^{(n)}) + R(x), \qquad \text{con tal que} (v', u^{(n-1)}) = R(x).$$

Supongamos, por el momento, que se pueden satisfacer las n condiciones impuestas a v. Sea $L = D^n + P_1(x)D^{n-1} + \ldots + P_n(x)$. Cuando aplicamos L a y_1 encontramos

$$\begin{aligned} L(y_1) &= y_1^{(n)} + P_1(x)y_1^{(n-1)} + \cdots + P_n(x)y_1 \\ &= \{(v, u^{(n)}) + R(x)\} + P_1(x)(v, u^{(n-1)}) + \cdots + P_n(x)(v, u) \\ &= (v, L(u)) + R(x) = (v, 0) + R(x) = R(x). \end{aligned}$$

Es decir, $L(y_1) = R(x)$, así que y_1 es una solución de la ecuación no homogénea.

El método tendrá éxito si se pueden satisfacer las n condiciones que hemos impuesto a v. Esas condiciones establecen que $(v', u^{(k)}) = 0$ para $k = 0, 1, \ldots, n - 2$, y que $(v', u^{(n-1)}) = R(x)$. Podemos escribir esas n ecuaciones en una sola ecuación matricial,

$$W(x)v'(x) = R(x)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{6.23}$$

donde $v(x)$ se considera como una matriz columna $n\times 1$, y W es la matriz $n\times n$ cuyas filas están formadas por los componentes de u y sus derivadas sucesivas:

$$W = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \\ u_1' & u_2' & \cdots & u_n' \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ u_1^{(n-1)} & u_2^{(n-1)} & \cdots & u_n^{(n-1)} \end{bmatrix}.$$

La matriz W se llama *matriz wronskiana* de $u_1, \ldots, u_n$, en recuerdo de J. M. H. Wronski (1778-1853).

En la sección que sigue demostraremos que la matriz wronskiana es no singular. Por consiguiente, podemos multiplicar ambos miembros de (6.23) por $W(x)^{-1}$ obteniendo

$$v'(x) = R(x)W(x)^{-1}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Elijamos dos puntos c y x en el intervalo J que se considera, e integremos esa ecuación vectorial en el intervalo, desde c a x, obteniendo

$$v(x) = v(c) + \int_c^x R(t)W(t)^{-1}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} dt = v(c) + z(x),$$

donde

$$z(x) = \int_c^x R(t)W(t)^{-1}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} dt.$$

La fórmula $y_1 = (u, v)$ para la solución particular toma ahora el aspecto

$$y_1 = (u, v) = (u, v(c) + z) = (u, v(c)) + (u, z).$$

El primer término $(u, v(c))$ satisface la ecuación homogénea puesto que es una combinación lineal de $u_1, \ldots, u_n$. Por consiguiente podemos omitir este término y utilizar el segundo término (u, z) como una solución particular de la ecuación no homogénea. Es decir, una solución particular de $L(y) = R$ viene dada por el producto interior

$$(u(x), z(x)) = \left(u(x), \int_c^x R(t)W(t)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} dt\right).$$

Obsérvese que no es necesario que la función R sea continua en el intervalo J. Todo lo que se exige es que R sea integrable en $[c, x]$.

Podemos resumir los resultados de esta sección mediante el teorema siguiente.

TEOREMA 6.11. *Sean $u_1, \ldots, u_n$, n soluciones independientes de la ecuación diferencial lineal de orden n homogénea $L(y) = 0$ en un intervalo J. Entonces una solución particular y_1 de la ecuación no homogénea $L(y) = R$ viene dada por la fórmula*

$$y_1(x) = \sum_{k=1}^{n} u_k(x)v_k(x),$$

donde $v_1, \ldots, v_n$ son los elementos de la matriz columna v, $n \times 1$, determinada por la ecuación

$$v(x) = \int_c^x R(t)W(t)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} dt. \tag{6.24}$$

En esta fórmula, W es la matriz wronskiana de $u_1, \ldots, u_n$, y c es un punto cualquiera de J.

Observación: La integral definida en (6.24) puede reemplazarse por cualquier integral indefinida.

$$\int R(x)W(x)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} dx.$$

EJEMPLO 1. Hallar la solución general de la ecuación diferencial

$$y'' - y = \frac{2}{1 + e^x}$$

en el intervalo $(-\infty, +\infty)$.

Solución. La ecuación homogénea, $(D^2 - 1)y = 0$ tiene las dos soluciones independientes $u_1(x) = e^x$, $u_2(x) = e^{-x}$. La matriz wronskiana de u_1 y u_2 es

$$W(x) = \begin{bmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{bmatrix}.$$

Puesto que $\det W(x) = -2$, la matriz es no singular y su inversa viene dada por

$$W(x)^{-1} = -\frac{1}{2}\begin{bmatrix} -e^{-x} & -e^{-x} \\ -e^x & e^x \end{bmatrix}.$$

Por consiguiente,

$$W(x)^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} -e^{-x} \\ e^x \end{bmatrix}$$

y tenemos

$$R(x)W(x)^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{1}{2}\frac{2}{1+e^x}\begin{bmatrix} -e^{-x} \\ e^x \end{bmatrix} = \begin{bmatrix} \dfrac{e^{-x}}{1+e^x} \\ \dfrac{-e^x}{1+e^x} \end{bmatrix}$$

Integrando cada componente del vector del segundo miembro encontramos

$$v_1(x) = \int \frac{e^{-x}}{1+e^x}\,dx = \int \left(e^{-x} - 1 + \frac{e^x}{1+e^x}\right) dx = -e^{-x} - x + \log(1+e^x)$$

y

$$v_2(x) = \int \frac{-e^x}{1+e^x}\,dx = -\log(1+e^x).$$

La solución general de la ecuación diferencial es, por consiguiente,

$$\begin{aligned} y &= c_1u_1(x) + c_2u_2(x) + v_1(x)u_1(x) + v_2(x)u_2(x) \\ &= c_1e^x + c_2e^{-x} - 1 - xe^x + (e^x - e^{-x})\log(1+e^x). \end{aligned}$$

6.12 No singularidad de la matriz wronskiana de n soluciones independientes de una ecuación lineal homogénea

En esta sección demostramos que la matriz wronskiana W de n soluciones independientes $u_1, \ldots, u_n$ de una ecuación homogénea $L(y) = 0$ es no singular. Lo vamos a hacer demostrando que el determinante de W es una función exponen-

cial que nunca se anula en el intervalo J que se considera.

Sea $w(x) = \det W(x)$ para cada x de J, y supongamos que la ecuación diferencial que se satisface con las funciones $u_1, \ldots, u_n$ tiene la forma

$$y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = 0. \tag{6.25}$$

Tenemos entonces:

TEOREMA 6.12. *El determinante wronskiano satisface la ecuación diferencial de primer orden*

$$w' + P_1(x)w = 0 \tag{6.26}$$

en J. Por consiguiente, si $c \in J$ tenemos

$$w(x) = w(c) \exp\left[-\int_c^x P_1(t)\,dt\right] \quad \textit{(Fórmula de Abel).} \tag{6.27}$$

Además, $w(x) \neq 0$ para todo x de J.

Demostración. Sea u el vector fila $u = (u_1, \ldots, u_n)$. Ya que cada componente de u satisface la ecuación diferencial (6.25) lo mismo ocurre para u. Las filas de la matriz wronskiana W son los vectores $u, u', \ldots, u^{(n-1)}$. Luego podemos escribir

$$w = \det W = \det(u, u', \ldots, u^{(n-1)}).$$

La derivada de w es el determinante de la matriz obtenida derivando la última fila de W (ver ejercicio 8 de la sección 3.17). Esto es

$$w' = \det(u, u', \ldots, u^{(n-2)}, u^{(n)}).$$

Multiplicando la última fila de w por $P_1(x)$ también tenemos

$$P_1(x)w = \det(u, u', \ldots, u^{(n-2)}, P_1(x)u^{(n-1)}).$$

Sumando esas dos últimas ecuaciones encontramos

$$w' + P_1(x)w = \det(u, u', \ldots, u^{(n-2)}, u^{(n)} + P_1(x)u^{(n-1)}).$$

Pero las filas de este último determinante son dependientes puesto que u satisface la ecuación diferencial (6.25). Por tanto el determinante es cero, lo que significa que w satisface (6.26). Resolviendo (6.26) obtenemos la fórmula de Abel (6.27).

Seguidamente demostramos que $w(c) \neq 0$ para un cierto c de J. Lo hacemos por reducción al absurdo. Supongamos que $w(t) = 0$ para todo t de J. Elijamos un t fijo en J, por ejemplo $t = t_0$, y consideremos el sistema lineal de ecuaciones algebraicas

$$W(t_0)X = O,$$

donde X es un vector columna. Ya que $\det W(t_0) = 0$, la matriz $W(t_0)$ es singular por lo que ese sistema admite una solución no nula, sea ésta $X = (c_1, \ldots, c_n) \neq (0, \ldots, 0)$. Utilizando los componentes de este vector no nulo, sea f la combinación lineal

$$f(t) = c_1u_1(t) + \cdots + c_nu_n(t).$$

La función f así definida satisface $L(f) = 0$ en J puesto que es una combinación lineal de $u_1, \ldots, u_n$. La ecuación matricial $W(t_0)X = O$ implica que

$$f(t_0) = f'(t_0) = \cdots = f^{(n-1)}(t_0) = 0.$$

Por consiguiente, f tiene como valor inicial el vector O en $t = t_0$, así que, según el teorema de unicidad, f es la solución cero. Esto significa $c_1 = \ldots = c_n = 0$, lo lo cual es una contradicción. Por lo tanto $w(t) \neq 0$ para un cierto t en J. Tomando ese valor de t como c en la fórmula de Abel vemos que $w(x) \neq 0$ para todo x de J. Esto completa la demostración del teorema 6.12.

6.13 Métodos especiales para determinar una solución particular de la ecuación no homogénea. Reducción a un sistema de ecuaciones lineales de primer orden

Si bien el método de variación de constantes proporciona un método general para determinar una solución particular de $L(y) = R$, existen métodos especiales que a menudo son de aplicación más sencilla cuando la ecuación tiene ciertas formas especiales. Por ejemplo, si la ecuación tiene coeficientes constantes podemos reducir el problema al de resolver una sucesión de ecuaciones lineales de primer orden. El método general queda mejor ilustrado con un ejemplo sencillo.

EJEMPLO 1. Hallar una solución particular de la ecuación

$$(D - 1)(D - 2)y = xe^{x+x^2}. \tag{6.28}$$

Solución. Pongamos $u = (D - 2)y$. Entonces la ecuación se transforma en

$$(D - 1)u = xe^{x+x^2}.$$

Ésta es una ecuación lineal de primer orden en u que puede resolverse con el teorema 6.1. Una solución particular es

$$u = \tfrac{1}{2}e^{x+x^2}.$$

Sustituyéndola en la ecuación $u = (D - 2)y$ obtenemos

$$(D - 2)y = \tfrac{1}{2}e^{x+x^2},$$

que es una ecuación de primer orden en y. Resolviéndola con el teorema 6.1 encontramos que una solución particular (con $y_1(0) = 0$) viene dada por

$$y_1(x) = \tfrac{1}{2}e^{2x}\int_0^x e^{t^2-t}\,dt.$$

Aunque la integral no puede calcularse por medio de funciones elementales, consideramos la ecuación como resuelta, puesto que la solución se expresa por medio de integrales de funciones corrientes. La solución general de (6.28) es

$$y = c_1e^x + c_2e^{2x} + \tfrac{1}{2}e^{2x}\int_0^x e^{t^2-t}\,dt.$$

6.14 Método del anulador para determinar una solución particular de la ecuación no homogénea

Describimos a continuación un método que se puede utilizar si la ecuación $L(y) = R$ tiene coeficientes constantes y si el segundo miembro R es anulado por un operador de coeficientes constantes, a saber $A(R) = 0$. En principio, el método es muy sencillo. Aplicamos el operador A a los dos miembros de la ecuación diferencial $L(y) = R$ y obtenemos una nueva ecuación $AL(y) = 0$ que debe ser satisfecha por todas las soluciones de la ecuación original. Puesto que AL es otro operador con coeficientes constantes podemos determinar su núcleo calculando las raíces de la ecuación característica de AL. El problema se reduce entonces a determinar, a partir de ese núcleo, una función particular y_1 que satisface $L(y_1) = R$. Veamos algunos ejemplos.

EJEMPLO 1. Hallar una solución particular de la ecuación

$$(D^4 - 16)y = x^4 + x + 1.$$

Solución. El segundo miembro, un polinomio de cuarto grado, es anulado por el operador D^5. Por tanto, cualquier solución de la ecuación dada también lo es de la ecuación

(6.29) $$D^5(D^4 - 16)y = 0.$$

Las raíces de la ecuación característica son 0, 0, 0, 0, 0, 2, -2, $2i$, $-2i$, con lo que todas las soluciones de (6.29) se encuentran en la combinación lineal

$$y = c_1 + c_2x + c_3x^2 + c_4x^3 + c_5x^4 + c_6e^{2x} + c_7e^{-2x} + c_8 \cos 2x + c_9 \operatorname{sen} 2x .$$

Deseamos elegir las c_i de modo que $L(y) = x^4 + x + 1$, donde $L = D^4 - 16$. Puesto que los cuatro últimos términos son anulados por L, podemos tomar $c_6 = c_7 = c_8 = c_9 = 0$ e intentamos encontrar $c_1, \ldots, c_5$ de modo que

$$L(c_1 + c_2x + c_3x^2 + c_4x^3 + c_5x^4) = x^4 + x + 1 .$$

Es decir, buscamos una solución particular y_1 que sea un polinomio de cuarto grado que satisfaga $L(y_1) = x^4 + x + 1$. Para simplificar escribamos

$$16y_1 = ax^4 + bx^3 + cx^2 + dx + e .$$

Esto nos da $16y_1^{(4)} = 24a$, así que $y_1^{(4)} = 3a/2$. Sustituyendo en la ecuación diferencial $L(y_1) = x^4 + x + 1$, tenemos que determinar a, b, c, d, e que satisfagan

$$\tfrac{3}{2}a - ax^4 - bx^3 - cx^2 - dx - e = x^4 + x + 1 .$$

Igualando los coeficientes de las potencias semejantes de x obtenemos

$$a = -1 , \qquad b = c = 0 , \qquad d = -1 , \qquad e = -\tfrac{5}{2} ,$$

de manera que la solución particular y_1 viene dada por

$$y_1 = -\tfrac{1}{16}x^4 - \tfrac{1}{16}x - \tfrac{5}{32} .$$

EJEMPLO 2. Resolver la ecuación diferencial $y'' - 5y' + 6y = xe^x$.

Solución. La ecuación diferencial tiene la forma

$$L(y) = R , \tag{6.30}$$

donde $R(x) = xe^x$ y $L = D^2 - 5D + 6$. La correspondiente ecuación homogénea puede ponerse en la forma

$$(D - 2)(D - 3)y = 0 ;$$

que tiene las soluciones independientes $u_1(x) = e^{2x}$, $u_2(x) = e^{3x}$.

Buscamos ahora una solución particular y_1 de la ecuación no homogénea. Comprobamos que la función $R(x) = xe^x$ es una solución de la ecuación homogénea

$$(D - 1)^2 y = 0 .$$

Por consiguiente, si aplicamos a los dos miembros de (6.30) el operador $(D - 1)^2$ encontramos que cualquier función que satisfaga (6.30) debe también satisfacer la ecuación

$$(D - 1)^2(D - 2)(D - 3)y = 0 .$$

Esta ecuación diferencial tiene las raíces características 1, 1, 2, 3, por lo que todas sus soluciones se encuentran en la combinación lineal

$$y = ae^x + bxe^x + ce^{2x} + de^{3x} ,$$

donde a, b, c, d son constantes. Queremos elegir a, b, c, d, de modo que la solución y_1 resultante satisfaga $L(y_1) = xe^x$. Puesto que $L(ce^{2x} + de^{3x}) = 0$ cualesquiera que sean los c y d, sólo necesitamos elegir a y b de manera que $L(ae^x + bxe^x) = xe^x$ y tomamos $c = d = 0$. Si ponemos

$$y_1 = ae^x + bxe^x ,$$

tenemos

$$D(y_1) = (a + b)e^x + bxe^x , \qquad D^2(y_1) = (a + 2b)e^x + bxe^x ,$$

así que la ecuación $(D^2 - 5D + 6)y_1 = xe^x$ se convierte en

$$(2a - 3b)e^x + 2bxe^x = xe^x .$$

Suprimiendo el factor e^x e igualando los coeficientes de las potencias semejantes de x encontramos $a = \frac{3}{4}, b = \frac{1}{2}$. Por lo tanto, $y_1 = \frac{3}{4}e^x + \frac{1}{2}xe^x$ y la solución general de $L(y) = R$ viene dada por la fórmula

$$y = c_1e^{2x} + c_2e^{3x} + \tfrac{3}{4}e^x + \tfrac{1}{2}xe^x .$$

El método utilizado en los ejemplos anteriores se llama *método del anulador.* Se podrá emplear si podemos encontrar un operador de coeficientes constantes A que anule R. A partir de nuestro estudio de las ecuaciones diferenciales lineales homogéneas con coeficientes constantes, sabemos que las únicas funciones de valo-

res reales anuladas por operadores con coeficientes constantes son combinaciones lineales de funciones de la forma

$$x^{m-1}e^{\alpha x}, \qquad x^{m-1}e^{\alpha x}\cos\beta x, \qquad x^{m-1}e^{\alpha x}\operatorname{sen}\beta x,$$

donde m es un entero positivo y α y β son constantes reales. La función $y = x^{m-1}e^{\alpha x}$ es una solución de una ecuación diferencial con raíz característica α de multiplicidad m. Por lo tanto, esta función tiene el anulador $(D - \alpha)^m$. Cada una de las funciones $y = x^{m-1}e^{\alpha x}\cos\beta x$ e $y = x^{m-1}e^{\alpha x}\operatorname{sen}\beta x$ es una solución de una ecuación diferencial con raíces características complejas $\alpha \pm i\beta$, con multiplicidad m cada una, de modo que son anuladas por el operador $[D^2 - 2\alpha D + (\alpha^2 + \beta^2)]^m$. Para facilitar la resolución de algunos ejercicios, damos una lista de esos anuladores en la tabla 6.1, junto con alguno de sus casos particulares.

TABLA 6.1

Función	*Anulador*
$y = x^{m-1}$	D^m
$y = e^{\alpha x}$	$D - \alpha$
$y = x^{m-1}e^{\alpha x}$	$(D - \alpha)^m$
$y = \cos\beta x$ o $y = \operatorname{sen}\beta x$	$D^2 + \beta^2$
$y = x^{m-1}\cos\beta x$ o $y = x^{m-1}\operatorname{sen}\beta x$	$(D^2 + \beta^2)^m$
$y = e^{\alpha x}\cos\beta x$ o $y = e^{\alpha x}\operatorname{sen}\beta x$	$D^2 - 2\alpha D + (\alpha^2 + \beta^2)$
$y = x^{m-1}e^{\alpha x}\cos\beta x$ o $y = x^{m-1}e^{\alpha x}\operatorname{sen}\beta x$	$[D^2 - 2\alpha D + (\alpha^2 + \beta^2)]^m$

Aunque el método del anulador es muy eficaz cuando se puede aplicar, se limita a las ecuaciones cuyos segundos miembros R tienen un anulador con coeficientes constantes. Si $R(x)$ tiene la forma e^{x^2}, $\log x$, o $\tan x$, el método no se podrá aplicar; tenemos entonces que utilizar la variación de constantes o algún otro método para encontrar una solución particular.

6.15 Ejercicios

En cada uno de los ejercicios 1 al 10, hallar la solución general en el intervalo $(-\infty, +\infty)$.

1. $y'' - y' = x^2$.
2. $y'' - 4y = e^{2x}$.
3. $y'' + 2y' = 3xe^x$.
4. $y'' + 4y = \operatorname{sen} x$.
5. $y'' - 2y' + y = e^x + e^{2x}$.
6. $y''' - y' = e^x$.
7. $y''' - y' = e^x + e^{-x}$.
8. $y''' + 3y'' + 3y' + y = xe^{-x}$.
9. $y'' + y = xe^x \operatorname{sen} 2x$.
10. $y^{(4)} - y = x^2e^{-x}$.

11. Si un operador con coeficientes constantes A anula f y si un operador con coeficientes constantes B anula g, demostrar que el producto AB anula $f + g$.
12. Sea A un operador con coeficientes constantes con polinomio característico p_A.
a) Aplicar el método del anulador para demostrar que la ecuación diferencial $A(y) = e^{\alpha x}$ tiene una solución particular de la forma

$$y_1 = \frac{e^{\alpha x}}{p_A(\alpha)}$$

si α no es un cero del polinomio p_A.
b) Si α es un cero simple de p_A (multiplicidad 1), demostrar que la ecuación $A(y) = e^{\alpha x}$ tiene la solución particular

$$y_1 = \frac{xe^{\alpha x}}{p'_A(\alpha)}.$$

c) Generalizar los resultados de a) y b) cuando α es un cero de p_A con multiplicidad m.
13. Dados dos operadores con coeficientes constantes A y B cuyos polinomios característicos no tienen ceros comunes. Sea $C = AB$.
a) Demostrar que toda solución de la ecuación diferencial $C(y) = 0$ tiene la forma $y = y_1 + y_2$, donde $A(y_1) = 0$ y $B(y_2) = 0$.
b) Demostrar que las funciones y_1 e y_2 de la parte a) están determinadas con unicidad. Esto es, para una y dada que satisfaga $C(y) = 0$ existe sólo un par y_1, y_2 con las propiedades de la parte a).
14. Si $L(y) = y'' + ay' + by$, donde a y b son constantes, sea f la solución particular de $L(y) = 0$ que satisface las condiciones $f(0) = 0$ y $f'(0) = 1$. Demostrar que una solución particular de $L(y) = R$ viene dada por la fórmula

$$y_1(x) = \int_c^x f(x - t)R(t)\, dt$$

para cualquier c. En particular, si las raíces de la ecuación característica son iguales, por ejemplo $r_1 = r_2 = m$, demostrar que la fórmula que da $y_1(x)$ se convierte en

$$y_1(x) = e^{mx} \int_c^x (x - t)e^{-mt}R(t)\, dt.$$

15. Sea Q el operador «multiplicación por x». Esto es, $Q(y)(x) = x \cdot y(x)$ para cada y de la clase $\mathscr{C}^\infty$ y cada x real. Designemos con I el operador identidad, definido por $I(y) = y$ para cada y de $\mathscr{C}^\infty$.
a) Demostrar que $DQ - QD = I$.
b) Demostrar que $D^2Q - QD^2$ es un operador de primer orden con coeficientes constantes, y determinar tal operador explícitamente como un polinomio de primer grado en D.
c) Demostrar que $D^3Q - QD^3$ es un operador de segundo orden con coeficientes constantes, y determinar dicho operador explícitamente como un polinomio de segundo grado en D.
d) Enunciar la generalización sugerida por el operador $D^nQ - QD^n$, y demostrarla por inducción.

En cada uno de los ejercicios del 16 al 20, hallar una solución general de la ecuación diferencial en el intervalo que se indica.

16. $y'' - y = 1/x, \qquad (0, +\infty).$
17. $y'' + 4y = \sec 2x, \qquad \left(-\frac{\pi}{4}, \frac{\pi}{4}\right).$
18. $y'' - y = \sec^3 x - \sec x, \qquad \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$
19. $y'' - 2y' + y = e^{e^x}(e^x - 1)^2, \qquad (-\infty, +\infty).$
20. $y''' - 7y'' + 14y' - 8y = \log x, \qquad (0, +\infty).$

6.16 Ejercicios varios sobre ecuaciones diferenciales lineales

1. Una curva integral $y = u(x)$ de la ecuación diferencial $y'' - 3y' - 4y = 0$ corta a una curva integral $y = v(x)$ de la ecuación diferencial $y'' + 4y' - 5y = 0$ en el origen. Determinar las funciones u y v si las dos curvas tienen la misma pendiente en el origen y si

$$\lim_{x\to\infty} \frac{[v(x)]^4}{u(x)} = \frac{5}{6}.$$

2. Una curva integral $y = u(x)$ de la ecuación diferencial $y'' - 4y' + 29y = 0$ corta a una curva integral $y = v(x)$ de la ecuación diferencial $y'' + 4y' + 13y = 0$ en el origen. Las dos curvas tienen la misma pendiente en el origen. Determinar u y v si $u'(\pi/2) = 1$.
3. Sabiendo que la ecuación diferencial $y'' + 4xy' + Q(x)y = 0$ tiene dos soluciones de la forma $y_1 = u(x)$ e $y_2 = xu(x)$, donde $u(0) = 1$, determinar $u(x)$ y $Q(x)$ explícitamente en función de x.
4. Sea $L(y) = y'' + P_1y' + P_2y$. Para resolver la ecuación no homogénea $L(y) = R$ por el método de variación de constantes, necesitamos conocer dos soluciones linealmente independientes de la ecuación homogénea. Este ejercicio demuestra que si se conoce *una* solución u_1 de $L(y) = 0$, y si u_1 no se anula en un intervalo J, una segunda solución u_2 de la ecuación homogénea viene dada por la fórmula

$$u_2(x) = u_1(x) \int_c^x \frac{Q(t)}{[u_1(t)]^2}\, dt,$$

donde $Q(x) = e^{-\int P_1(x)dx}$, y c es un punto cualquiera de J. Estas dos soluciones son independientes en J.
a) Demostrar que la función u_2 satisface $L(y) = 0$.
b) Demostrar que u_1 y u_2 son independientes en J.
5. Hallar la solución general de la ecuación

$$xy'' - 2(x + 1)y' + (x + 2)y = x^3e^{2x}$$

para $x > 0$, sabiendo que la ecuación homogénea tiene una solución de la forma $y = e^{mx}$.

6. Obtener a simple inspección una solución no nula y luego encontrar la solución general de la ecuación diferencial

$$(y'' - 4y') + x^2(y' - 4y) = 0.$$

7. Hallar la solución general de la ecuación diferencial

$$4x^2y'' + 4xy' - y = 0,$$

sabiendo que existe una solución particular de la forma $y = x^m$ para $x > 0$.

8. Hallar una solución de la ecuación homogénea por tanteo, y encontrar luego la solución general de la ecuación

$$x(1 - x)y'' - (1 - 2x)y' + (x^2 - 3x + 1)y = (1 - x)^3.$$

9. Hallar la solución general de la ecuación

$$(2x - 3x^3)y'' + 4y' + 6xy = 0,$$

sabiendo que tiene una solución que es un polinomio en x.

10. Hallar la solución general de la ecuación

$$x^2(1 - x)y'' + 2x(2 - x)y' + 2(1 + x)y = x^2,$$

sabiendo que la ecuación homogénea tiene una solución de la forma $y = x^c$.

11. Sea $g(x) = \int_1^x e^t/t\,dt$ si $x > 0$. (No intentar el cálculo de esta integral.) Hallar todos los valores de la constante a tales que la función f definida por

$$f(x) = \frac{1}{x}e^{ag(x)}$$

satisfaga la ecuación diferencial lineal

$$x^2y'' + (3x - x^2)y' + (1 - x - e^{2x})y = 0.$$

Sabiendo esto determinar la solución general de la ecuación en el intervalo $(0, +\infty)$.

6.17 Ecuaciones lineales de segundo orden con coeficientes analíticos

Se dice que una función f es analítica en un intervalo $(x_0 - r, x_0 + r)$ si f

tiene un desarrollo en serie de potencias en ese intervalo,

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n ,$$

convergente para $|x - x_0| < r$. Si los coeficientes de una ecuación diferencial lineal homogénea

$$y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = 0$$

son analíticos en un intervalo $(x_0 - r, x_0 + r)$, puede entonces demostrarse que existen n soluciones independientes $u_1, \ldots, u_n$, cada una de las cuales es analítica en el mismo intervalo. Demostraremos este teorema para ecuaciones de segundo orden y luego discutiremos un importante ejemplo que se presenta en muchas aplicaciones.

TEOREMA 6.13. *Sean P_1 y P_2 analíticas en un intervalo abierto $(x_0 - r, x_0 + r)$, tales como*

$$P_1(x) = \sum_{n=0}^{\infty} b_n(x - x_0)^n , \qquad P_2(x) = \sum_{n=0}^{\infty} c_n(x - x_0)^n .$$

Entonces la ecuación diferencial

$$y'' + P_1(x)y' + P_2(x)y = 0 \tag{6.31}$$

tiene dos soluciones independientes u_1 y u_2 que son analíticas en el mismo intervalo.

Demostración. Intentemos encontrar una serie de potencias solución de la forma

$$y = \sum_{n=0}^{\infty} a_n(x - x_0)^n , \tag{6.32}$$

convergente en el intervalo dado. Para ello, sustituimos las series dadas para P_1 y P_2 en la ecuación diferencial y determinamos luego las relaciones que los coeficientes a_n deben satisfacer para que la función y dada por (6.32) satisfaga la ecuación.

Las derivadas y' e y'' pueden obtenerse derivando la serie de potencias de y término a término (ver teorema 11.9 del Volumen I). Esto nos da

$$y' = \sum_{n=1}^{\infty} na_n(x - x_0)^{n-1} = \sum_{n=0}^{\infty}(n + 1)a_{n+1}(x - x_0)^n,$$

$$y'' = \sum_{n=2}^{\infty} n(n - 1)a_n(x - x_0)^{n-2} = \sum_{n=0}^{\infty}(n + 2)(n + 1)a_{n+2}(x - x_0)^n.$$

Los productos $P_1(x)y'$ y $P_2(x)y$ vienen dados por las series de potencias (*)

$$P_1(x)y' = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}(k + 1)a_{k+1}b_{n-k}\right)(x - x_0)^n$$

y

$$P_2(x)y = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a_kc_{n-k}\right)(x - x_0)^n.$$

Cuando esas series se sustituyen en la ecuación diferencial (6.31) encontramos

$$\sum_{n=0}^{\infty}\left\{(n + 2)(n + 1)a_{n+2} + \sum_{k=0}^{n}[(k + 1)a_{k+1}b_{n-k} + a_kc_{n-k}]\right\}(x - x_0)^n = 0.$$

Por consiguiente, la ecuación diferencial se satisfará si elegimos los coeficientes a_n de modo que satisfagan las fórmulas recurrentes

$$(6.33) \qquad (n + 2)(n + 1)a_{n+2} = -\sum_{k=0}^{n}[(k + 1)a_{k+1}b_{n-k} + a_kc_{n-k}]$$

para $n = 0, 1, 2, \ldots$. Esta fórmula expresa a_{n+2} en función de los coeficientes anteriores $a_0, a_1, \ldots, a_{n+1}$ y los coeficientes de las funciones dadas P_1 y P_2. Elijamos valores arbitrarios de los dos primeros coeficientes a_0 y a_1 y utilicemos la fórmula recurrente para definir los restantes coeficientes $a_2, a_3, \ldots$, en función de a_0 y a_1. Esto garantiza que la serie de potencias (6.32) satisfará la ecuación diferencial (6.31). El siguiente paso de la demostración es probar que la serie así definida converge realmente para todo x del intervalo $(x_0 - r, x_0 + r)$. Esto se hace mayorando la serie (6.32) por otra serie de potencias convergente. Finalmente, demostramos que podemos elegir a_0 y a_1 para obtener dos soluciones independientes.

Demostremos ahora que la serie (6.32) cuyos coeficientes están definidos por (6.33) converge en el intervalo indicado.

(*) Los lectores no familiarizados con la multiplicación de series de potencias pueden consultar el ejercicio 7 de la sección 6.21.

Elijamos un punto fijo, $x_1 \neq x_0$ en el intervalo $(x_0 - r,\ x_0 + r)$ y sea $t = |x_1 - x_0|$. Puesto que las series que definen P_1 y P_2 convergen absolutamente para $x = x_1$ los términos de esas series están acotados, pongamos

$$|b_k|\, t^k \le M_1 \qquad \text{y} \qquad |c_k|\, t^k \le M_2,$$

para ciertos números $M_1 > 0$, $M_2 > 0$. Sea M el mayor de los dos números M_1 y tM_2. Tenemos entonces

$$|b_k| \le \frac{M}{t^k} \qquad \text{y} \qquad |c_k| \le \frac{M}{t^{k+1}}.$$

La fórmula de recurrencia implica la desigualdad

$$\begin{aligned}(n + 2)(n + 1)\, |a_{n+2}| &\le \sum_{k=0}^{n}\left\{(k + 1)\, |a_{k+1}|\, \frac{M}{t^{n-k}} + |a_k|\, \frac{M}{t^{n-k+1}}\right\}\\ &= \frac{M}{t^{n+1}}\left\{\sum_{k=0}^{n}(k + 1)\, |a_{k+1}|\, t^{k+1} + \sum_{k=0}^{n}|a_{k+1}|\, t^{k+1} + |a_0| - |a_{n+1}|\, t^{n+1}\right\}\\ &\le \frac{M}{t^{n+1}}\left\{\sum_{k=0}^{n}(k + 2)\, |a_{k+1}|\, t^{k+1} + |a_0|\right\} = \frac{M}{t^{n+1}} \sum_{k=0}^{n+1}(k + 1)\, |a_k|\, t^k.\end{aligned}$$

Hagamos ahora $A_0 = |a_0|$, $A_1 = |a_1|$, y definamos $A_2, A_3, \ldots$ sucesivamente mediante la fórmula de recurrencia

$$(6.34) \qquad (n + 2)(n + 1)A_{n+2} = \frac{M}{t^{n+1}} \sum_{k=0}^{n+1}(k + 1)A_k t^k$$

para $n \ge 0$. Entonces $|a_n| \le A_n$ para todo $n \ge 0$, así que la serie $\sum a_n(x - x_0)^n$ es mayorada por la serie $\sum A_n|x - x_0|^n$. Aplicamos ahora el criterio del cociente para demostrar que $\sum A_n|x - x_0|^n$ converge si $|x - x_0| < t$.

Sustituyendo n por $n - 1$ en (6.34) y restando el producto de la ecuación por t^{-1} de (6.34) encontramos que $(n + 2)(n + 1)A_{n+2} - t^{-1}(n + 1)nA_{n+1} = M(n + 2)A_{n+1}$. Por consiguiente

$$A_{n+2} = A_{n+1}\, \frac{(n + 1)n + (n + 2)Mt}{(n + 2)(n + 1)t},$$

y hallamos

$$\frac{A_{n+2}\, |x - x_0|^{n+2}}{A_{n+1}\, |x - x_0|^{n+1}} = \frac{(n + 1)n + (n + 2)Mt}{(n + 2)(n + 1)t}\, |x - x_0| \to \frac{|x - x_0|}{t}$$

cuando $n \to \infty$. Este límite es menor que 1 si $|x - x_0| < t$. Luego $\sum a_n(x - x_0)^n$ converge si $|x - x_0| < t$. Pero ya que $t = |x_1 - x_0|$ y que x_1 es un punto arbitrario en el intervalo $(x_0 - r, x_0 + r)$, la serie $\sum a_n(x - x_0)^n$ converge para todo x en $(x_0 - r, x_0 + r)$.

Los dos primeros coeficientes a_0 y a_1 representan los valores iniciales de y y de su derivada en el punto x_0. Si u_1 es la serie de potencias solución con $a_0 = 1$ y $a_1 = 0$, de manera que

$$u_1(x_0) = 1 \qquad \text{y} \qquad u_1'(x_0) = 0,$$

y u_2 la solución con $a_0 = 0$ y $a_1 = 1$, de tal modo que

$$u_2(x_0) = 0 \qquad \text{y} \qquad u_2'(x_0) = 1,$$

entonces las soluciones u_1 y u_2 serán independientes. Esto completa la demostración.

6.18 La ecuación de Legendre

En esta sección encontramos series de potencias soluciones para la ecuación de Legendre,

$$(1 - x^2)y'' - 2xy' + \alpha(\alpha + 1)y = 0, \tag{6.35}$$

donde α es cualquier constante real. Esta ecuación se presenta en problemas de atracción y de flujo de calor con simetría esférica. Cuando α es un entero positivo encontraremos que la ecuación tiene soluciones polinómicas llamadas *polinomios de Legendre*. Esos son los mismos polinomios que se encontraron al estudiar el proceso de Gram-Schmidt (capítulo 1, pág. 31).

La ecuación de Legendre puede escribirse así:

$$[(x^2 - 1)y']' = \alpha(\alpha + 1)y,$$

que tiene la forma

$$T(y) = \lambda y,$$

en donde T es un operador de Sturm-Liouville, $T(f) = (pf')'$, con $p(x) = x^2 - 1$ y $\lambda = \alpha(\alpha + 1)$. Por consiguiente, las soluciones no nulas de la ecuación de Legendre son autofunciones de T correspondientes al autovalor $\alpha(\alpha + 1)$. Puesto que $p(x)$ satisface las condiciones de contorno

$$p(1) = p(-1) = 0,$$

el operador T es simétrico con respecto al producto interior

$$(f, g) = \int_{-1}^{1} f(x)g(x)\, dx .$$

La teoría general de los operadores simétricos nos dice que las autofunciones correspondientes a autovalores distintos son ortogonales (teorema 5.3).

En la ecuación diferencial tratada en el teorema 6.13 el coeficiente de y'' es 1. La ecuación de Legendre puede ponerse en esa forma si dividimos por $1 - x^2$. De (6.35) obtenemos

$$y'' + P_1(x)y' + P_2(x)y = 0 ,$$

donde

$$P_1(x) = -\frac{2x}{1 - x^2} \qquad \text{y} \qquad P_2(x) = \frac{\alpha(\alpha + 1)}{1 - x^2} ,$$

si $x^2 \neq 1$. Puesto que $1/(1 - x^2) = \sum_{n=0}^{\infty} x^{2n}$ para $|x| < 1$, P_1 y P_2 tienen ambos desarrollos en serie de potencias en el intervalo abierto $(-1, 1)$ así que el teorema 6.13 es aplicable. Para encontrar la fórmula de recurrencia para los coeficientes es más sencillo dejar la ecuación en la forma (6.35) e intentar hallar una serie de potencias solución de la forma

$$y = \sum_{n=0}^{\infty} a_n x^n$$

válida en el intervalo abierto $(-1, 1)$. Derivando esta serie término a término obtenemos

$$y' = \sum_{n=1}^{\infty} n a_n x^{n-1} \qquad \text{y} \qquad y'' = \sum_{n=2}^{\infty} n(n - 1)a_n x^{n-2} .$$

Por lo tanto, tenemos

$$2xy' = \sum_{n=1}^{\infty} 2n a_n x^n = \sum_{n=0}^{\infty} 2n a_n x^n ,$$

y

$$\begin{aligned}(1 - x^2)y'' &= \sum_{n=2}^{\infty} n(n - 1)a_n x^{n-2} - \sum_{n=2}^{\infty} n(n - 1)a_n x^n \\ &= \sum_{n=0}^{\infty} (n + 2)(n + 1)a_{n+2} x^n - \sum_{n=0}^{\infty} n(n - 1)a_n x^n \\ &= \sum_{n=0}^{\infty} [(n + 2)(n + 1)a_{n+2} - n(n - 1)a_n]x^n .\end{aligned}$$

Si sustituimos esas series en la ecuación diferencial (6.35), vemos que la ecuación se satisface si, y sólo si, los coeficientes satisfacen la relación

$$(n+2)(n+1)a_{n+2} - n(n-1)a_n - 2na_n + \alpha(\alpha+1)a_n = 0$$

para todo $n \geq 0$. Esta ecuación es la misma que

$$(n+2)(n+1)a_{n+2} - (n-\alpha)(n+1+\alpha)a_n = 0,$$

o

(6.36) $$a_{n+2} = -\frac{(\alpha-n)(\alpha+n+1)}{(n+1)(n+2)}\,a_n.$$

Esta relación nos permite determinar $a_2, a_4, a_6, \ldots$, sucesivamente en función de a_0. Análogamente, podemos calcular $a_3, a_5, a_7, \ldots$, en función de a_1. Para los coeficientes con índices pares tenemos

$$a_2 = -\frac{\alpha(\alpha+1)}{1\cdot 2}\,a_0,$$

$$a_4 = -\frac{(\alpha-2)(\alpha+3)}{3\cdot 4}\,a_2 = (-1)^2\frac{\alpha(\alpha-2)(\alpha+1)(\alpha+3)}{4!}\,a_0,$$

y, en general,

$$a_{2n} = (-1)^n\frac{\alpha(\alpha-2)\cdots(\alpha-2n+2)\cdot(\alpha+1)(\alpha+3)\cdots(\alpha+2n-1)}{(2n)!}\,a_0.$$

Esto puede demostrarse por inducción. Para los coeficientes con índices impares encontramos

$$a_{2n+1} = (-1)^n\frac{(\alpha-1)(\alpha-3)\cdots(\alpha-2n+1)\cdot(\alpha+2)(\alpha+4)\cdots(\alpha+2n)}{(2n+1)!}\,a_1.$$

Por consiguiente, la serie que define y puede escribirse así

(6.37) $$y = a_0u_1(x) + a_1u_2(x),$$

donde

(6.38)
$$u_1(x) = 1 + \sum_{n=1}^{\infty}(-1)^n\frac{\alpha(\alpha-2)\cdots(\alpha-2n+2)\cdot(\alpha+1)(\alpha+3)\cdots(\alpha+2n-1)}{(2n)!}\,x^{2n}$$

y

(6.39)

$$u_2(x) = x + \sum_{n=1}^{\infty}(-1)^n \frac{(\alpha - 1)(\alpha - 3)\cdots(\alpha - 2n + 1)\cdot(\alpha + 2)(\alpha + 4)\cdots(\alpha + 2n)}{(2n + 1)!} x^{2n+1}.$$

El criterio del cociente demuestra que cada una de esas series converge para $|x| < 1$. Asimismo, ya que la relación (6.36) se satisface separadamente por los coeficientes pares e impares, cada una de las funciones u_1 y u_2 es una solución de la ecuación diferencial (6.35). Esas soluciones satisfacen las condiciones iniciales

$$u_1(0) = 1, \qquad u_1'(0) = 0, \qquad u_2(0) = 0, \qquad u_2'(0) = 1.$$

Puesto que u_1 y u_2 son independientes, la solución general de la ecuación de Legendre (6.35) en el intervalo abierto $(-1, 1)$ viene dada por la combinación lineal (6.37) con las constantes arbitrarias a_0 y a_1.

Cuando α es 0 o un entero par positivo, por ejemplo $\alpha = 2m$, la serie de $u_1(x)$ se transforma en un polinomio de grado $2m$ que sólo contiene potencias pares de x. Puesto que tenemos

$$\alpha(\alpha - 2)\cdots(\alpha - 2n + 2) = 2m(2m - 2)\cdots(2m - 2n + 2) = \frac{2^n m!}{(m - n)!}$$

y

$$(\alpha + 1)(\alpha + 3)\cdots(\alpha + 2n - 1) = (2m + 1)(2m + 3)\cdots(2m + 2n - 1)$$
$$= \frac{(2m + 2n)!\, m!}{2^n (2m)!\,(m + n)!}$$

la fórmula que da $u_1(x)$ en este caso se transforma en

(6.40)
$$u_1(x) = 1 + \frac{(m!)^2}{(2m!)} \sum_{k=1}^{m}(-1)^k \frac{(2m + 2k)!}{(m - k)!\,(m + k)!\,(2k)!} x^{2k}.$$

Por ejemplo, cuando $\alpha = 0, 2, 4, 6$ $(m = 0, 1, 2, 3)$ los polinomios correspondientes son

$$u_1(x) = 1, \qquad 1 - 3x^3, \qquad 1 - 10x^2 + \tfrac{35}{3}x^4, \qquad 1 - 21x^2 + 63x^4 - \tfrac{231}{5}x^6.$$

La serie que da $u_2(x)$ no es un polinomio cuando α es par debido a que el coeficiente de x^{2n+1} nunca es cero.

Cuando α es un entero *impar*, se invierten los papeles de u_1 y u_2; la serie

correspondiente a $u_2(x)$ se convierte en un polinomio y la serie de $u_1(x)$ no es un polinomio. Concretamente, si $\alpha = 2m + 1$, tenemos

$$(6.41)\quad u_2(x) = x + \frac{(m!)^2}{(2m+1)!}\sum_{k=1}^{m}(-1)^k \frac{(2m+2k+1)!}{(m-k)!\,(m+k)!\,(2k+1)!}\,x^{2k+1}.$$

Por ejemplo, cuando $\alpha=1, 3, 5$ $(m=0, 1, 2)$, los polinomios correspondientes son

$$u_2(x) = x, \qquad x - \tfrac{5}{3}x^3, \qquad x - \tfrac{14}{3}x^3 + \tfrac{21}{5}x^5.$$

6.19 Polinomios de Legendre

Algunas de las propiedades de las soluciones polinómicas de la ecuación de Legendre pueden deducirse directamente a partir de la ecuación diferencial o de las fórmulas (6.40) y (6.41). Otras se deducen más fácilmente partiendo de otra fórmula para esos polinomios que deduciremos ahora.

Obtendremos primero una única fórmula que contiene (salvo factores constantes) los polinomios (6.40) y (6.41). Pongamos

$$(6.42)\qquad P_n(x) = \frac{1}{2^n}\sum_{r=0}^{[n/2]}\frac{(-1)^r(2n-2r)!}{r!\,(n-r)!\,(n-2r)!}\,x^{n-2r},$$

donde $[n/2]$ representa el mayor entero $\leq n/2$. Demostraremos luego que ése es el *polinomio de Legendre* de grado n introducido en el capítulo 1. Cuando n es par, es un múltiplo constante del polinomio $u_1(x)$ de la ecuación (6.40); cuando n es impar, es un múltiplo constante del polinomio $u_2(x)$ de (6.41). (*) Los siete primeros polinomios de Legendre vienen dados por las fórmulas

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \tfrac{1}{2}(3x^2 - 1), \qquad P_3(x) = \tfrac{1}{2}(5x^3 - 3x),$$

$$P_4(x) = \tfrac{1}{8}(35x^4 - 30x^2 + 3), \qquad P_5(x) = \tfrac{1}{8}(63x^5 - 70x^3 + 15x),$$

$$P_6(x) = \tfrac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5).$$

La figura 6.1 muestra las gráficas de las cinco primeras de esas funciones en el intervalo $[-1, 1]$.

(*) Cuando n es par, $n = 2m$, podemos reemplazar el índice de sumación k en (6.40) por otro r, siendo $r = m - k$; encontramos que la suma (6.40) es el producto de $P_n(x)$ por una constante. Análogamente, cuando n es impar. un cambio de índice transforma la suma (6.41) en el producto de $P_n(x)$ por una constante.

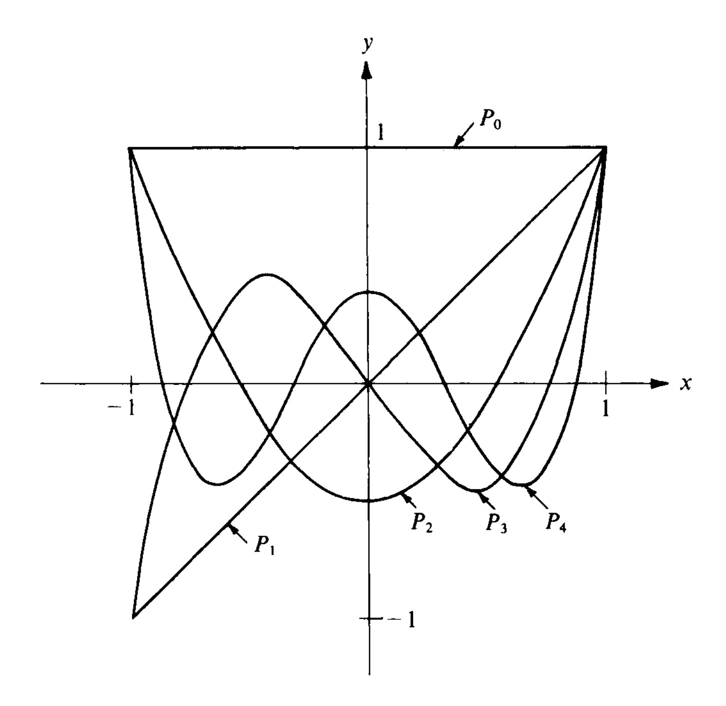


FIGURA 6.1 *Gráficas de los polinomios de Legendre en el intervalo* $[-1, 1]$

Podemos ahora demostrar que, salvo factores escalares, los polinomios de Legendre son los que se obtienen aplicando el proceso de ortogonalización de Gram-Schmidt a la sucesión de polinomios $1, x, x^2, \ldots,$ con el producto interior

$$(f, g) = \int_{-1}^{1} f(x)g(x)\, dx\,.$$

Observemos primero que si $m \neq n$ los polinomios P_n y P_m son ortogonales porque son autofunciones de un operador simétrico pertenecientes a autovalores distintos. También, puesto que P_n tiene grado n y $P_0 = 1$, los polinomios $P_0(x), P_1(x), \ldots, P_n(x)$ engendran el mismo subespacio que $1, x, \ldots, x^n$. En la sección 1.14, ejemplo 2, construimos otro conjunto ortogonal de polinomios $y_0, y_1, y_2, \ldots,$ tales que $y_0(x), y_1(x), \ldots, y_n(x)$ engendran el mismo subespacio que $1, x, \ldots, x^n$ para cada n. El teorema de ortogonalización (teorema 1.13) nos dice que, salvo factores escalares, existe un solo conjunto de funciones ortogonales con esa propiedad. Luego debe ser

$$P_n(x) = c_n y_n(x)$$

para ciertos escalares c_n. El coeficiente de x^n en $y_n(x)$ es 1, así que c_n es el coeficiente de x^n en $P_n(x)$. De (6.42) vemos que

$$c_n = \frac{(2n)!}{2^n(n!)^2}.$$

6.20 Fórmula de Rodrigues para los polinomios de Legendre

En la suma (6.42) que define $P_n(x)$ observamos que

$$\frac{(2n-2r)!}{(n-2r)!}x^{n-2r} = \frac{d^n}{dx^n}x^{2n-2r} \qquad \text{y} \qquad \frac{1}{r!\,(n-r)!} = \frac{1}{n!}\binom{n}{r},$$

donde $\binom{n}{r}$ es el coeficiente binomial, y escribimos la suma en la forma

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\sum_{r=0}^{[n/2]}(-1)^r\binom{n}{r}x^{2n-2r}.$$

Cuando $[n/2] < r \leq n$, el término x^{2n-2r} tiene grado menor que n, por lo que su derivada *n-ésima* es cero. Por consiguiente, no se altera la suma si hacemos que r recorra los valores de 0 a n. Esto nos da

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\sum_{r=0}^{n}(-1)^r\binom{n}{r}x^{2n-2r}.$$

Vemos ahora que la suma del segundo miembro es el desarrollo del binomio $(x^2-1)^n$. Por consiguiente, tenemos

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n.$$

Ésta se conoce con el nombre de *fórmula de Rodrigues*, en honor de Olinde Rodrigues (1794-1851), economista y reformador francés.

Con la fórmula de Rodrigues y la ecuación diferencial, podemos deducir un número de propiedades importantes de los polinomios de Legendre. Algunas de esas propiedades se citan a continuación. Sus demostraciones están esbozadas en el conjunto de ejercicios que sigue.

Para cada $n \geq 0$ tenemos

$$P_n(1) = 1.$$

Además, $P_n(x)$ es el único polinomio que satisface la ecuación de Legendre

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0$$

y vale 1 cuando $x = 1$.

Para cada $n \geq 0$ tenemos

$$P_n(-x) = (-1)^n P_n(x).$$

Esto demuestra que P_n es una función par cuando n es par, y una función impar si n es impar.

Ya hemos mencionado la relación de ortogonalidad,

$$\int_{-1}^{1} P_n(x)P_m(x)\,dx = 0 \qquad \text{si} \quad m \neq n.$$

Cuando $m = n$ tenemos la expresión de la norma

$$\|P_n\|^2 = \int_{-1}^{1} [P_n(x)]^2\,dx = \frac{2}{2n + 1}.$$

Todo polinomio de grado n puede expresarse como combinación lineal de polinomios de Legendre $P_0, P, \ldots, P_n$. En efecto, si f es un polinomio de grado n tenemos

$$f(x) = \sum_{k=0}^{n} c_k P_k(x),$$

donde

$$c_k = \frac{2k + 1}{2} \int_{-1}^{1} f(x)P_k(x)\,dx.$$

A partir de la relación de ortogonalidad resulta que

$$\int_{-1}^{1} g(x)P_n(x)\,dx = 0$$

para todo polinomio g de grado menor que n. Esta propiedad puede usarse para demostrar que el polinomio de Legendre P_n tiene n ceros reales y distintos situados todos en el intervalo $(-1, 1)$.

6.21 Ejercicios

1. La ecuación de Legendre (6.35) con $\alpha = 0$ tiene el polinomio solución $u_1(x) = 1$ y una

solución no polinómica, dada por la serie (6.41).
a) Demostrar que la suma de la serie que da u_2 viene dada por

$$u_2(x) = \frac{1}{2}\log\frac{1+x}{1-x} \quad \text{para } |x| < 1 .$$

b) Comprobar directamente que la función u_2 de la parte a) es una solución de la ecuación de Legendre cuando $\alpha = 0$.

2. Demostrar que la función f definida por la ecuación

$$f(x) = 1 - \frac{x}{2}\log\frac{1+x}{1-x}$$

para $|x| < 1$ satisface la ecuación de Legendre (6.35) con $\alpha = 1$. Expresar esta función como combinación lineal de las soluciones u_1 y u_2 dadas por las ecuaciones (6.38) y (6.39).

3. La ecuación de Legendre (6.35) puede escribirse en la forma

$$[(x^2 - 1)y']' - \alpha(\alpha + 1)y = 0 .$$

a) Si a, b, c son constantes siendo $a > b$ y $4c + 1 > 0$, demostrar que la ecuación diferencial del tipo

$$[(x - a)(x - b)y']' - cy = 0$$

puede transformarse en una ecuación de Legendre mediante un cambio de variable de la forma $x = At + B$, con $A > 0$. Determinar A y B en función de a y b.
b) Utilizar el método sugerido en la parte a) para transformar la ecuación

$$(x^2 - x)y'' + (2x - 1)y' - 2y = 0$$

en una ecuación de Legendre.

4. Hallar dos soluciones independientes, en forma de serie de potencias, de la *ecuación de Hermite*

$$y'' - 2xy' + 2\alpha y = 0$$

en un intervalo de la forma $(-r, r)$. Demostrar que una de esas soluciones es un polinomio cuando α es entero no negativo.

5. Hallar una serie de potencias solución de la ecuación diferencial

$$xy'' + (3 + x^3)y' + 3x^2y = 0$$

válida para todo x. Hallar una segunda solución de la forma $y = x^{-2}\sum a_n x^n$ válida para $x \neq 0$.

6. Hallar una serie de potencias solución de la ecuación diferencial

$$x^2y'' + x^2y' - (\alpha x + 2)y = 0$$

válida en un intervalo de la forma $(-r, r)$.

7. Dadas dos funciones A y B analíticas en un intervalo $(x_0 - r, x_0 + r)$,

$$A(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n, \qquad B(x) = \sum_{n=0}^{\infty} b_n(x - x_0)^n.$$

Puede demostrarse que el producto $C(x) = A(x)B(x)$ es también analítica en $(x_0 - r, x_0 + r)$. Este ejercicio demuestra que C tiene el desarrollo en serie de potencias

$$C(x) = \sum_{n=0}^{\infty} c_n(x - x_0)^n, \qquad \text{donde} \quad c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

a) Con la regla de Leibniz para la derivada n-ésima de un producto demostrar que la derivada n-ésima de C viene dada por

$$C^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} A^{(k)}(x) B^{(n-k)}(x).$$

b) Hacer uso del hecho de que $A^{(k)}(x_0) = k!\, a_k$ y $B^{(n-k)}(x_0) = (n - k)!\, b_{n-k}$ para obtener

$$C^{(n)}(x_0) = n! \sum_{k=0}^{n} a_k b_{n-k}.$$

Puesto que $C^{(n)}(x_0) = n!\, c_n$, queda demostrada la fórmula que da c_n.

En los ejercicios del 8 al 14, $P_n(x)$ designa el polinomio de Legendre de grado n. Esos ejercicios dan el esquema de las demostraciones de las propiedades de los polinominos de Legendre descritas en la sección 6.20.

8. a) Con la fórmula de Rodrigues demostrar que

$$P_n(x) = \frac{1}{2^n}(x + 1)^n + (x - 1)Q_n(x),$$

donde $Q_n(x)$ es un polinomio.

b) Demostrar que $P_n(1) = 1$ y que $P_n(-1) = (-1)^n$.

c) Demostrar que $P_n(x)$ es el único polinomio solución de la ecuación de Legendre (con $\alpha = n$) que tiene valor 1 cuando $x = 1$.

9. a) Utilizar las ecuaciones diferenciales a las que satisfacen P_n y P_m para demostrar que

$$[(1 - x^2)(P_n P_m' - P_n' P_m)]' = [n(n + 1) - m(m + 1)]P_n P_m.$$

b) Si $m \neq n$, integrar la ecuación del apartado a) desde -1 a 1 para dar otra demostración de la relación de ortogonalidad

$$\int_{-1}^{1} P_n(x) P_m(x)\, dx = 0.$$

10. a) Sea $f(x) = (x^2 - 1)^n$. Con la integración por partes demostrar que

$$\int_{-1}^{1} f^{(n)}(x) f^{(n)}(x)\, dx = -\int_{-1}^{1} f^{(n+1)}(x) f^{(n-1)}(x)\, dx .$$

Aplicar esta fórmula reiteradamente para deducir que la integral del primer miembro es igual a

$$2(2n)! \int_{0}^{1} (1 - x^2)^n\, dx .$$

b) La sustitución $x = \cos t$ transforma la integral $\int_0^1 (1 - x^2)^n\, dx$ en $\int_0^{\pi/2} \operatorname{sen}^{2n+1} t\, dt$. Emplear la relación

$$\int_{0}^{\pi/2} \sin^{2n+1} t\, dt = \frac{2n(2n-2)\cdots 2}{(2n+1)(2n-1)\cdots 3 \cdot 1}$$

y la fórmula de Rodrigues para obtener

$$\int_{-1}^{1} [P_n(x)]^2\, dx = \frac{2}{2n+1} .$$

11. a) Demostrar que

$$P_n(x) = \frac{(2n)!}{2^n (n!)^2} x^n + Q_n(x),$$

donde $Q_n(x)$ es un polinomio de grado menor que n.
b) Expresar el polinomio $f(x) = x^4$ como combinación lineal de P_0, P_1, P_2, P_3, y P_4.
c) Demostrar que todo polinomio f de grado n puede expresarse como combinación lineal de los polinomios de Legendre $P_0, P_1, \ldots, P_n$.

12. a) Si f es un polinomio de grado n, escribir

$$f(x) = \sum_{k=0}^{n} c_k P_k(x) .$$

[Esto es posible según el ejercicio 11 c).] Para un m fijo, $0 \le m \le n$, multiplicar ambos miembros de esa ecuación por $P_m(x)$ e integrar entre -1 y 1. Teniendo en cuenta los ejercicios 9 b) y 10 b) deducir la relación

$$c_m = \frac{2m+1}{2} \int_{-1}^{1} f(x) P_m(x)\, dx .$$

13. Utilizar los ejercicios 9 y 11 para demostrar que $\int_{-1}^{1} g(x) P_n(x)\, dx = 0$ para todo polinomio g de grado menor que n.

14. a) Con el auxilio del teorema de Rolle demostrar que P_n no puede tener ceros múltiples en el intervalo abierto $(-1, 1)$. Es decir, los ceros de P_n que están en $(-1, 1)$ deben ser ceros simples.
b) Supongamos que P_n tiene m ceros en el intervalo $(-1, 1)$. Si $m = 0$, sea $Q_0(x) = 1$. Si $m \geq 1$, sea

$$Q_m(x) = (x - x_1)(x - x_2) \cdots (x - x_m),$$

donde $x_1, x_2, \ldots, x_m$ son m ceros de P_n en $(-1, 1)$. Demostrar que, en cada punto x de $(-1, 1)$, $Q_m(x)$ tiene el mismo signo que $P_n(x)$.
c) Con el apartado b) y el ejercicio 13, demostrar que la desigualdad $m < n$ nos lleva a una contradicción. Esto demuestra que P_n tiene n ceros reales distintos, todos situados en el intervalo abierto $(-1,1)$.
15. a) Demostrar que el valor de la integral $\int_{-1}^{1} P_n(x)P'_{n+1}(x)\,dx$ es independiente de n.
b) Calcular la integral $\int_{-1}^{1} x\,P_n(x)P_{n-1}(x)\,dx$.

6.22 Método de Frobenius

En la sección 6.17 aprendimos a encontrar soluciones, en forma de series de potencias, de la ecuación diferencial

$$(6.43) \qquad y'' + P_1(x)y' + P_2(x)y = 0$$

en un entorno del punto x_0 en el que los coeficientes P_1 y P_2 son analíticos. Si uno de los dos P_1 o P_2 no es analítico en las proximidades de x_0, pueden o no existir series de potencias solución en el entorno de x_0. Por ejemplo, supongamos que intentamos encontrar una serie de potencias solución de la ecuación diferencial

$$(6.44) \qquad x^2y'' - y' - y = 0$$

en un entorno de $x_0 = 0$. Si suponemos que existe una solución $y = \sum a_k x^k$ y la sustituimos en la ecuación diferencial, llegamos a la fórmula de recurrencia

$$a_{n+1} = \frac{n^2 - n - 1}{n + 1}\, a_n .$$

Si bien esto nos da una serie de potencias $y = \sum a_k x^k$ que formalmente satisface (6.44), el criterio del cociente pone en evidencia que esa serie converge *únicamente* para $x = 0$. Así pues, no existe serie de potencias solución de (6.44) válida en cualquier intervalo abierto en torno a $x_0 = 0$. Este ejemplo no contradice el teorema 6.13 porque cuando ponemos la ecuación (6.44) en la forma (6.43) encontramos que los coeficientes P_1 y P_2 vienen dados por

$$P_1(x) = -\frac{1}{x^2} \qquad \text{y} \qquad P_2(x) = -\frac{1}{x^2}.$$

Estas funciones no admiten desarrollo en serie de potencias en torno al origen. La dificultad aquí radica en que el coeficiente de y'' en (6.44) tiene el valor 0 cuando $x = 0$; es decir, la ecuación diferencial tiene un punto singular en $x = 0$.

Para apreciar las dificultades que se presentan en la investigación de las ecuaciones diferenciales en las proximidades de un punto singular es necesario conocer la teoría de funciones de variable compleja. Sin embargo, algunos casos particulares de ecuaciones con puntos singulares pueden tratarse con métodos elementales. Por ejemplo, supongamos la ecuación diferencial (6.43) equivalente a una ecuación de la forma

$$(x - x_0)^2y'' + (x - x_0)P(x)y' + Q(x)y = 0, \tag{6.45}$$

donde P y Q tienen desarrollos en series de potencias en un cierto intervalo abierto $(x_0 - r, x_0 + r)$. En este caso decimos que x_0 es un punto singular *regular* de la ecuación. Si dividimos ambos miembros de (6.45) por $(x - x_0)^2$ la ecuación se transforma en

$$y'' + \frac{P(x)}{x - x_0}y' + \frac{Q(x)}{(x - x_0)^2}y = 0$$

para $x \neq x_0$. Si $P(x_0) \neq 0$ o $Q(x_0) \neq 0$, o si $Q(x_0) = 0$ y $Q'(x_0) \neq 0$, o el coeficiente de y' o el de y no tendrá desarrollo en serie de potencias en torno al punto x_0, así que el teorema 6.13 no es aplicable. En 1873 el matemático alemán Jorge Frobenius (1849-1917) desarrolló un método muy práctico para tratar tales ecuaciones. Enunciaremos el teorema de Frobenius pero no daremos su demostración. (*) En la sección siguiente damos los detalles de la demostración para un caso particular importante, la ecuación de Bessel.

El teorema de Frobenius se desdobla en dos partes, que dependen de la naturaleza de las raíces de la ecuación cuadrática

$$t(t - 1) + P(x_0)t + Q(x_0) = 0. \tag{6.46}$$

Esta ecuación cuadrática se llama la *ecuación de índices* de la ecuación diferencial dada (6.45). Los coeficientes $P(x_0)$ y $Q(x_0)$ son los términos constantes de los desarrollos de P y Q en series de potencias. Sean α_1 y α_2 las raíces de la ecuación de índices. El tipo de solución que se obtiene por el método de Frobenius depende de si esas raíces difieren o no en un número entero.

(*) Para la demostración véase E. Hille, *Analysis,* Vol. II, Blaisdell Publishing Co., 1966, o E. A. Coddington, *An Introduction to Ordinary Differential Equations,* Prentice-Hall, 1961.

TEOREMA 6.14. PRIMER CASO DEL TEOREMA DE FROBENIUS. *Sean α_1 y α_2 las raíces de la ecuación de índices y supongamos que $\alpha_1 - \alpha_2$ no es un entero. Entonces la ecuación diferencial* (6.45) *tiene dos soluciones independientes u_1 y u_2 de la forma*

$$u_1(x) = |x - x_0|^{\alpha_1} \sum_{n=0}^{\infty} a_n (x - x_0)^n, \qquad \text{con} \quad a_0 = 1\,; \tag{6.47}$$

y

$$u_2(x) = |x - x_0|^{\alpha_2} \sum_{n=0}^{\infty} b_n (x - x_0)^n, \qquad \text{con} \quad b_0 = 1\,. \tag{6.48}$$

Ambas series convergen en el intervalo $|x - x_0| < r$, y la ecuación diferencial se satisface para $0 < |x - x_0| < r$.

TEOREMA 6.15. SEGUNDO CASO DEL TEOREMA DE FROBENIUS. *Sean α_1 y α_2 las raíces de la ecuación de índices y supongamos que $\alpha_1 - \alpha_2 = N$ es un entero no negativo. Entonces la ecuación diferencial* (6.45) *tiene una solución u_1 de la forma* (6.47) *y otra solución independiente u_2 de la forma*

$$u_2(x) = |x - x_0|^{\alpha_2} \sum_{n=0}^{\infty} b_n (x - x_0)^n + C\, u_1(x) \log |x - x_0|\,,$$

donde $b_0 = 1$. La constante C no es nula si $N = 0$. Si $N > 0$, la constante C puede o no ser cero. Como en el caso 1, ambas series convergen en el intervalo $|x - x_0| < r$, y las soluciones son válidas para $0 < |x - x_0| < r$.

6.23 Ecuación de Bessel

En esta sección utilizamos el método sugerido por Frobenius para resolver la ecuación de Bessel

$$x^2 y'' + xy' + (x^2 - \alpha^2)y = 0\,, \tag{6.49}$$

donde α es una constante no negativa. Esta ecuación se emplea en problemas relativos a vibraciones de membranas, flujo de calor en cilindros, y propagación de corrientes eléctricas en conductores cilíndricos. Algunas de sus soluciones se conocen con el nombre de *funciones de Bessel*. También se presentan esas funciones en Teoría de Números. La ecuación tomó el nombre del astrónomo alemán F. W. Bessel (1784-1846), si bien ya aparece en las investigaciones de Daniel Bernoulli (1732) y de Euler (1764).

La ecuación de Bessel tiene la forma (6.45) con $x_0 = 0$, $P(x) = 1$, y

$Q(x) = x^2 - \alpha^2$, de modo que el punto x_0 es un punto singular regular. Puesto que P y Q son analíticas en toda recta real, intentamos encontrar soluciones de la forma

(6.50) $$y = |x|^t \sum_{n=0}^{\infty} a_n x^n ,$$

con $a_0 \neq 0$, válida para todo x real con la posible excepción de $x = 0$.

Primero mantenemos $x > 0$, de modo que $x^{t} = x^t$. La derivación de (6.50) nos da

$$y' = tx^{t-1} \sum_{n=0}^{\infty} a_n x^n + x^t \sum_{n=0}^{\infty} n a_n x^{n-1} = x^{t-1} \sum_{n=0}^{\infty} (n + t) a_n x^n .$$

Análogamente, obtenemos

$$y'' = x^{t-2} \sum_{n=0}^{\infty} (n + t)(n + t - 1) a_n x^n .$$

Si $L(y) = x^2 y'' + xy' + (x^2 - \alpha^2)y$, encontramos

$$L(y) = x^t \sum_{n=0}^{\infty} (n + t)(n + t - 1) a_n x^n + x^t \sum_{n=0}^{\infty} (n + t) a_n x^n$$
$$+ x^t \sum_{n=0}^{\infty} a_n x^{n+2} - x^t \sum_{n=0}^{\infty} \alpha^2 a_n x^n = x^t \sum_{n=0}^{\infty} [(n + t)^2 - \alpha^2] a_n x^n + x^t \sum_{n=0}^{\infty} a_n x^{n+2} .$$

Pongamos ahora $L(y) = 0$, suprimamos x^t, e intentemos determinar los a_n de modo que el coeficiente de cada potencia de x se anule. Para el término independiente es necesario que $(t^2 - \alpha^2)a_0 = 0$. Como quiera que buscamos una solución con $a_0 \neq 0$, esto exige que

(6.51) $$t^2 - \alpha^2 = 0 .$$

Ésta es la *ecuación de índices.* Sus raíces α y $-\alpha$ son los únicos valores posibles de t que nos pueden dar una solución del tipo deseado.

Consideremos primero $t = \alpha$. Para esta t, las ecuaciones restantes para la determinación de los coeficientes se transforman en

(6.52) $$[(1 + \alpha)^2 - \alpha^2]a_1 = 0 \qquad \text{y} \qquad [(n + \alpha)^2 - \alpha^2]a_n + a_{n-2} = 0$$

para $n \geq 2$. Puesto que $\alpha \geq 0$, la primera de esas implica que $a_1 = 0$. La segunda

fórmula puede escribirse en la forma

(6.53) $$a_n = -\frac{a_{n-2}}{(n+\alpha)^2 - \alpha^2} = -\frac{a_{n-2}}{n(n+2\alpha)},$$

con lo que $a_3 = a_5 = a_7 = \ldots = 0$. Para los coeficientes con índices pares tenemos

$$a_2 = \frac{-a_0}{2(2+2\alpha)} = \frac{-a_0}{2^2(1+\alpha)}, \qquad a_4 = \frac{-a_2}{4(4+2\alpha)} = \frac{(-1)^2 a_0}{2^4 2!\,(1+\alpha)(2+\alpha)},$$

$$a_6 = \frac{-a_4}{6(6+2\alpha)} = \frac{(-1)^3 a_0}{2^6 3!\,(1+\alpha)(2+\alpha)(3+\alpha)},$$

y, en general

$$a_{2n} = \frac{(-1)^n a_0}{2^{2n} n!\,(1+\alpha)(2+\alpha)\cdots(n+\alpha)}.$$

Por consiguiente, la elección $t = \alpha$ nos da la solución

$$y = a_0 x^\alpha \left(1 + \sum_{n=1}^{\infty} \frac{(-1)^n x^{2n}}{2^{2n} n!\,(1+\alpha)(2+\alpha)\cdots(n+\alpha)}\right).$$

El criterio del cociente demuestra que la serie de potencias que aparece en esta fórmula converge para todo x real.

En esta discusión hemos supuesto que $x > 0$. Si $x < 0$ podemos repetirla reemplazando x^t por $(-x)^t$. Encontramos nuevamente que t debe satisfacer la ecuación $t^2 - \alpha^2 = 0$. Haciendo $t = \alpha$ obtenemos la misma solución, salvo que el factor externo x^α queda reemplazado por $(-x)^\alpha$. Por tanto la función f_α dada por la serie

(6.54) $$f_\alpha(x) = a_0 |x|^\alpha \left(1 + \sum_{n=1}^{\infty} \frac{(-1)^n x^{2n}}{2^{2n} n!\,(1+\alpha)(2+\alpha)\cdots(n+\alpha)}\right)$$

es una solución de la ecuación de Bessel válida para todo real $x \neq 0$. Para los valores de α para los que existen $f_\alpha'(0)$ y $f_\alpha''(0)$, la solución también es válida para $x = 0$.

Consideremos ahora la raíz $t = -\alpha$ de la ecuación de índices. Obtenemos en lugar de (6.52), las ecuaciones

$$[(1-\alpha)^2 - \alpha^2]a_1 = 0 \qquad \text{y} \qquad [(n-\alpha)^2 - \alpha^2]a_n + a_{n-2} = 0,$$

que se convierten en

$$(1 - 2\alpha)a_1 = 0 \qquad \text{y} \qquad n(n - 2\alpha)a_n + a_{n-2} = 0.$$

Si 2α no es entero esas ecuaciones nos dan $a_1 = 0$ y

$$a_n = -\frac{a_{n-2}}{n(n - 2\alpha)}$$

para n $\geq$ 2. Puesto que esta fórmula de recurrencia es la misma que (6.53), con α sustituida por $-\alpha$, llegamos a la solución

(6.55) $$f_{-\alpha}(x) = a_0 \, |x|^{-\alpha} \left(1 + \sum_{n=1}^{\infty} \frac{(-1)^n x^{2n}}{2^{2n} n! \, (1 - \alpha)(2 - \alpha) \cdots (n - \alpha)}\right)$$

válida para todo real $x \neq 0$.

La solución $f_{-\alpha}$ se obtuvo en la hipótesis de que 2α no es un entero positivo. No obstante, la serie que da $f_{-\alpha}$ tiene sentido incluso si 2α es un entero positivo, en tanto que α no sea entero positivo. Puede comprobarse que $f_{-\alpha}$ satisface la ecuación de Bessel para tal valor α. Por consiguiente, para cada $\alpha \geq 0$ tenemos la serie solución f_α, dada por (6.54); y si α no es un entero no negativo hemos encontrado otra solución $f_{-\alpha}$ dada por (6.55). Las dos soluciones f_α y $f_{-\alpha}$ son independientes, ya que una de ellas $\to \infty$ cuando $x \to 0$, y la otra no. Seguidamente simplificaremos la forma de las soluciones. Para ello necesitamos algunas propiedades de la función gamma de Euler que brevemente vamos a recordar.

Para cada real $s > 0$ definimos $\Gamma(s)$ mediante la integral impropia

$$\Gamma(s) = \int_{0+}^{\infty} t^{s-1} e^{-t} \, dt.$$

Esta integral converge si $s > 0$ y diverge si $s \leq 0$. La integración por partes nos conduce a la ecuación funcional

(6.56) $$\Gamma(s + 1) = s \, \Gamma(s).$$

Esto implica que

$$\Gamma(s + 2) = (s + 1)\Gamma(s + 1) = (s + 1)s \, \Gamma(s),$$

$$\Gamma(s + 3) = (s + 2)\Gamma(s + 2) = (s + 2)(s + 1)s \, \Gamma(s),$$

y, en general

(6.57) $$\Gamma(s + n) = (s + n - 1) \cdots (s + 1)s \, \Gamma(s)$$

para todo entero positivo n. Puesto que $\Gamma(1)= \int_0^\infty e^{-t}\,dt = 1$, cuando ponemos $s = 1$ en (6.57) encontramos

$$\Gamma(n + 1) = n!\,.$$

Así pues, la función gamma es una extensión de la función factorial de los enteros a los números reales positivos.

La ecuación funcional (6.56) puede usarse para extender la definición de $\Gamma(s)$ a valores negativos de s que no sean enteros. Escribamos (6.56) en la forma

$$\Gamma(s) = \frac{\Gamma(s + 1)}{s}\,. \tag{6.58}$$

El segundo miembro tiene sentido si $s + 1 > 0$ y $s \neq 0$. Por consiguiente, podemos usar esa ecuación para *definir* $\Gamma(s)$ si $-1 < s < 0$. El segundo miembro de (6.58) tiene ahora sentido si $s + 2 > 0$, $s \neq -1$, $s \neq 0$, y podemos usar esa ecuación para definir $\Gamma(s)$ para $-2 < s < -1$. Continuando así, podemos extender la definición de $\Gamma(s)$ por inducción a cualquier intervalo abierto de la forma $-n < s < -n + 1$, donde n es un entero positivo. La ecuación funcional (6.56) y su extensión (6.57) son ahora válidas para todo real s para el que ambos miembros tengan sentido.

Volvamos ahora a la discusión de la ecuación de Bessel. La serie que da f_α en la ecuación (6.54) contiene el producto $(1 + \alpha)(2 + \alpha)\ldots(n + \alpha)$. Podemos expresar este producto mediante la función gamma tomando $s = 1 + \alpha$ en (6.57). Esto nos da

$$(1 + \alpha)(2 + \alpha)\cdots(n + \alpha) = \frac{\Gamma(n + 1 + \alpha)}{\Gamma(1 + \alpha)}\,.$$

Por consiguiente, si elegimos $a_0 = 2^{-\alpha}/\Gamma(1 + \alpha)$ en (6.54) y designamos la función resultante $f_\alpha(x)$ por $J_\alpha(x)$ cuando $x > 0$, la solución para $x > 0$ puede escribirse en la forma

$$J_\alpha(x) = \left(\frac{x}{2}\right)^\alpha \sum_{n=0}^\infty \frac{(-1)^n}{n!\,\Gamma(n + 1 + \alpha)}\left(\frac{x}{2}\right)^{2n}. \tag{6.59}$$

La función J_α definida por esta ecuación para $x > 0$ y $\alpha \geq 0$ se llama *función de Bessel de primera especie y de orden* α. Cuando α es un entero no negativo, sea $\alpha = p$, la función de Bessel J_p viene dada por la serie de potencias

$$J_p(x) = \sum_{n=0}^\infty \frac{(-1)^n}{n!\,(n + p)!}\left(\frac{x}{2}\right)^{2n+p} \qquad (p = 0, 1, 2, \ldots).$$

Ésta es también solución de la ecuación de Bessel para $x < 0$. Se han construido tablas de funciones de Bessel. En la figura 6.2 se han dibujado las gráficas de las dos funciones de Bessel J_0 y J_1.

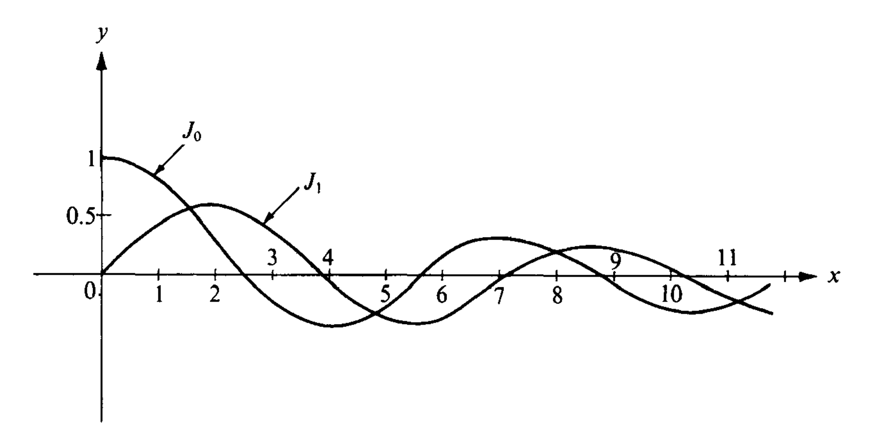


FIGURA 6.2 *Gráficas de las funciones de Bessel J_0 y J_1.*

Podemos definir una nueva función $J_{-\alpha}$ reemplazando α por $-\alpha$ en la ecuación (6.59), si α es tal que $\Gamma(n + 1 - \alpha)$ tiene sentido; esto es, si α no es un entero positivo. Por consiguiente, si $x > 0$ y $\alpha > 0$, $\alpha \neq 1, 2, 3, \ldots$, definimos

$$J_{-\alpha}(x) = \left(\frac{x}{2}\right)^{-\alpha} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\,\Gamma(n + 1 - \alpha)} \left(\frac{x}{2}\right)^{2n}.$$

Tomando $s = 1 - \alpha$ en (6.57) obtenemos

$$\Gamma(n + 1 - \alpha) = (1 - \alpha)(2 - \alpha) \cdots (n - \alpha)\,\Gamma(1 - \alpha)$$

y vemos que la serie que da $J_{-\alpha}(x)$ es la misma que la de $f_{-\alpha}(x)$ de la ecuación (6.55) con $a_0 = 2^{\alpha}/\Gamma(1 - \alpha)$, $x > 0$. Por lo tanto, si α no es entero positivo, $J_{-\alpha}$ es una solución de la ecuación de Bessel para $x > 0$.

Si α no es entero, las dos soluciones $J_\alpha(x)$ y $J_{-\alpha}(x)$ son linealmente independientes en el eje real positivo (puesto que su cociente no es constante) y la solución general de la ecuación de Bessel para $x > 0$ es

$$y = c_1 J_\alpha(x) + c_2 J_{-\alpha}(x).$$

Si α es un entero no negativo, sea $\alpha = p$, hemos encontrado únicamente la solución J_p y sus productos por constantes válidos para $x > 0$. Otra solución,

independiente de ésta, puede encontrarse por el método descrito en el ejercicio 4 de la sección 6.16. Éste establece que si u_1 es una solución de $y''+P_1y'+P_2y=0$ que nunca se anula en un intervalo I, una segunda solución u_2 independiente de u_1 viene dada por la integral

$$u_2(x) = u_1(x) \int_c^x \frac{Q(t)}{[u_1(t)]^2}\, dt,$$

donde $Q(x) = e^{-\int P_1(x)dx}$. Para la ecuación de Bessel tenemos $P_1(x) = 1/x$, así que $Q(x) = 1/x$ y una segunda solución u_2 viene dada por la fórmula

$$u_2(x) = J_p(x) \int_c^x \frac{1}{t[J_p(t)]^2}\, dt, \tag{6.60}$$

si c y x pertenecen al intervalo I en el que J_p no se anula.

Esta segunda solución puede ponerse en otras formas. Por ejemplo, a partir de (6.59) podemos escribir

$$\frac{1}{[J_p(t)]^2} = \frac{1}{t^{2p}}\, g_p(t),$$

donde $g_p(0) \neq 0$. En el intervalo I la función g_p tiene un desarrollo en serie de potencias

$$g_p(t) = \sum_{n=0}^{\infty} A_n t^n$$

que podría determinarse igualando coeficientes en la identidad $g_p(t)[J_p(t)]^2 = t^{2p}$. Si suponemos la existencia de tal desarrollo, el integrando de (6.60) toma la forma

$$\frac{1}{t[J_p(t)]^2} = \frac{1}{t^{2p+1}} \sum_{n=0}^{\infty} A_n t^n .$$

Integrando esta fórmula término a término entre c y x obtenemos un término logarítmico $A_{2p} \log x$ (de la potencia t^{-1}) más una serie de la forma $x^{-2p} \sum B_n x^n$. Por consiguiente, (6.60) toma la forma

$$u_2(x) = A_{2p} J_p(x) \log x + J_p(x) x^{-2p} \sum_{n=0}^{\infty} B_n x^n .$$

Puede demostrarse que el coeficiente $A_{2p} \neq 0$. Si multiplicamos $u_2(x)$ por $1/A_{2p}$ la solución que resulta se designa con $K_p(x)$ y tiene la forma

$$K_p(x) = J_p(x) \log x + x^{-p} \sum_{n=0}^{\infty} C_n x^n .$$

Esta es la forma de la solución anunciada en el segundo caso del teorema de Frobenius.

Habiendo llegado a esta fórmula, podemos comprobar que existe realmente una solución de esta forma sustituyendo en el segundo miembro de la ecuación de Bessel y determinando los coeficientes C_n de modo que la ecuación se satisfaga. Los detalles de este cálculo son largos y serán omitidos. El resultado final puede expresarse así

$$K_p(x) = J_p(x)\log x - \frac{1}{2}\left(\frac{x}{2}\right)^{-p}\sum_{n=0}^{p-1}\frac{(p-n-1)!}{n!}\left(\frac{x}{2}\right)^{2n} - \frac{1}{2}\left(\frac{x}{2}\right)^{p}\sum_{n=0}^{\infty}(-1)^n\frac{h_n + h_{n+p}}{n!\,(n+p)!}\left(\frac{x}{2}\right)^{2n},$$

donde $h_0 = 0$ y $h_n = 1 + \frac{1}{2} + \ldots + 1/n$ para $n \geq 1$. La serie del segundo miembro converge para todo real x. La función K_p definida para $x>0$ por esa fórmula se llama *función de Bessel de segunda especie y de orden p.* Puesto que K_p no es el producto de una constante por J_p, la solución general de la ecuación de Bessel en este caso para $x > 0$ es

$$y = c_1J_p(x) + c_2K_p(x).$$

En los ejercicios que siguen se discutirán otras propiedades de las funciones de Bessel.

6.24 Ejercicios

1. a) Sean f una solución cualquiera de la ecuación de Bessel de orden α y $g(x) = x^{1/2}f(x)$ para $x > 0$. Demostrar que g satisface la ecuación diferencial

$$y'' + \left(1 + \frac{1 - 4\alpha^2}{4x^2}\right)y = 0.$$

b) Cuando $4\alpha^2 = 1$ la ecuación diferencial del apartado a) se transforma en $y'' + y = 0$; su solución general es $y = A\cos x + B\,\text{sen}\,x$. Utilizar esa información y la igualdad (*) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ para demostrar que, para $x > 0$,

$$J_{1/2}(x) = \left(\frac{2}{\pi x}\right)^{1/2}\text{sen}\,x \qquad \text{y} \qquad J_{-1/2}(x) = \left(\frac{2}{\pi x}\right)^{1/2}\cos x.$$

(*) El cambio de variable $t = u^2$ nos da

$$\Gamma(\tfrac{1}{2}) = \int_{0+}^{\infty} t^{-1/2}e^{-t}\,dt = 2\int_0^{\infty} e^{-u^2}\,du = \sqrt{\pi}.$$

(En el ejercicio 16 de la sección 11.28 puede verse una demostración de que $2\int_0^{\infty} e^{-u^2}\,du = \sqrt{\pi}$.)

c) A partir de las series que dan $J_{1/2}(x)$ y $J_{-1/2}(x)$, deducir directamente las fórmulas del apartado b).

2. Utilizar la representación en series de potencias de las funciones de Bessel para demostrar que

a) $\dfrac{d}{dx}(x^{\alpha}J_{\alpha}(x)) = x^{\alpha}J_{\alpha-1}(x)$,

b) $\dfrac{d}{dx}(x^{-\alpha}J_{\alpha}(x)) = -x^{-\alpha}J_{\alpha+1}(x)$.

3. Sean $F_{\alpha}(x) = x^{\alpha}J_{\alpha}(x)$ y $G_{\alpha}(x) = x^{-\alpha}J_{\alpha}(x)$ para $x > 0$. Obsérvese que cada cero positivo de J_{α} es un cero de F_{α} y también lo es de G_{α}. Con el teorema de Rolle y el ejercicio 2 demostrar que los ceros de J_{α} y $J_{\alpha+1}$ se intercalan. Esto es, existe un cero de J_{α} entre cada par de ceros positivos de $J_{\alpha+1}$, y un cero de $J_{\alpha+1}$ entre cada par de ceros positivos de J_{α}. (Ver figura 6.2.)

4. a) A partir de las relaciones del ejercicio 2 deducir las relaciones de recurrencia

$$\frac{\alpha}{x}J_{\alpha}(x) + J'_{\alpha}(x) = J_{\alpha-1}(x) \qquad \text{y} \qquad \frac{\alpha}{x}J_{\alpha}(x) - J'_{\alpha}(x) = J_{\alpha+1}(x).$$

b) Con las relaciones del apartado a) deducir las fórmulas

$$J_{\alpha-1}(x) + J_{\alpha+1}(x) = \frac{2\alpha}{x}J_{\alpha}(x) \qquad \text{y} \qquad J_{\alpha-1}(x) - J_{\alpha+1}(x) = 2J'_{\alpha}(x).$$

5. Utilizar el ejercicio 1 b) y una fórmula de recurrencia adecuada para demostrar que

$$J_{3/2}(x) = \left(\frac{2}{\pi x}\right)^{1/2}\left(\frac{\operatorname{sen} x}{x} - \cos x\right).$$

Hallar una fórmula análoga para $J_{-3/2}(x)$. *Observación:* $J_{\alpha}(x)$ es una función elemental para todo α que sea la mitad de un entero impar.

6. Demostrar que

$$\frac{1}{2}\frac{d}{dx}(J_{\alpha}^2(x) + J_{\alpha+1}^2(x)) = \frac{\alpha}{x}J_{\alpha}^2(x) - \frac{\alpha+1}{x}J_{\alpha+1}^2(x)$$

y

$$\frac{d}{dx}(xJ_{\alpha}(x)J_{\alpha+1}(x)) = x(J_{\alpha}^2(x) - J_{\alpha+1}^2(x)).$$

7. a) Usar las identidades del ejercicio 6 para demostrar que

$$J_0^2(x) + 2\sum_{n=1}^{\infty}J_n^2(x) = 1 \qquad \text{y} \qquad \sum_{n=0}^{\infty}(2n+1)J_n(x)J_{n+1}(x) = \tfrac{1}{2}x.$$

b) Del apartado a) deducir que $|J_0(x)| \leq 1$ y $|J_n(x)| \leq \frac{1}{2}\sqrt{2}$ para $n = 1, 2, 3, \ldots$, y todo $x \geq 0$.

8. Sea $g_\alpha(x) = x^{1/2}f_\alpha(ax^b)$ para $x > 0$, donde a y b son constantes no nulas. Demostrar que g_α satisface la ecuación diferencial

$$x^2y'' + (a^2b^2x^{2b} + \tfrac{1}{4} - \alpha^2b^2)y = 0$$

si, y sólo si, f_α es una solución de la ecuación de Bessel de orden α.

9. Hacer uso del ejercicio 8 para expresar la solución general de cada una de las ecuaciones diferenciales siguientes por medio de las funciones de Bessel para $x > 0$.

a) $y'' + xy = 0$.
b) $y'' + x^2y = 0$.
c) $y'' + x^my = 0$.
d) $x^2y'' + (x^4 + \frac{1}{8})y = 0$.

10. Generalizar el ejercicio 8 cuando f_α y g_α se relacionan mediante la ecuación $g_\alpha(x) = x^cf_\alpha(ax^b)$ para $x > 0$. Hallar entonces la solución general de las siguientes ecuaciones por medio de las funciones de Bessel para $x > 0$.

a) $xy'' + 6y' + y = 0$.
b) $xy'' + 6y' + xy = 0$.
c) $xy'' + 6y' + x^4y = 0$.
d) $x^2y'' - xy' + (x + 1)y = 0$.

11. Dada la identidad de la forma

$$J_2(x) - J_0(x) = aJ_c''(x),$$

donde a y c son constantes. Determinar a y c.

12. Encontrar una serie de potencias solución de la ecuación diferencial $xy'' + y' + y = 0$ convergente para $-\infty < x < +\infty$. Demostrar que para $x > 0$ puede expresarse mediante una función de Bessel.

13. Consideremos una ecuación diferencial lineal de segundo orden de la forma

$$x^2A(x)y'' + xP(x)y' + Q(x)y = 0,$$

en la que $A(x)$, $P(x)$, y $Q(x)$ admiten desarrollos en series de potencias,

$$A(x) = \sum_{k=0}^{\infty} a_kx^k, \qquad P(x) = \sum_{k=0}^{\infty} p_kx^k, \qquad Q(x) = \sum_{k=0}^{\infty} q_kx^k,$$

con $a_0 \neq 0$, convergente cada una en un intervalo abierto $(-r, r)$. Si la ecuación diferencial tiene una serie solución de la forma

$$y = x^t\sum_{n=0}^{\infty} c_nx^n,$$

válida para $0 < x < r$, demostrar que t satisface una ecuación cuadrática de la forma $t^2 + bt + c = 0$, y determinar b y c en función de los coeficientes de las series de $A(x)$, $P(x)$ y $Q(x)$.

14. Considerar un caso particular del ejercicio 13 en el que $A(x) = 1 - x$, $P(x) = \frac{1}{2}$ y $Q(x) = -\frac{1}{2}x$. Hallar una serie solución con t no entero.

15. La ecuación diferencial $2x^2y'' + (x^2 - x)y' + y = 0$ tiene dos soluciones independientes de la forma

$$y = x^t\sum_{n=0}^{\infty} c_nx^n,$$

válida para $x > 0$. Determinar esas soluciones.

16. La ecuación diferencial no lineal $y'' + y + \alpha y^2 = 0$ es «ligeramente» no lineal si α es una constante pequeña no nula. Supongamos que existe una solución que puede expresarse como serie de potencias de α de la forma

$$y = \sum_{n=0}^{\infty} u_n(x)\alpha^n \quad \text{(válido en un cierto intervalo } 0<\alpha<r)$$

y que esta solución satisfaga las condiciones iniciales $y = 1$ e $y' = 0$ cuando $x = 0$. Para ajustarse a esas condiciones iniciales, intentamos elegir los coeficientes $u_n(x)$ de modo que $u_0(0) = 1$, $u_0'(0) = 0$ y $u_n(0) = u_n'(0) = 0$ para $n \geq 1$. Sustituir esa serie en la ecuación diferencial, igualar las potencias adecuadas de α y determinar con ello $u_0(x)$ y $u_1(x)$.

7

SISTEMAS DE ECUACIONES DIFERENCIALES

7.1 Introducción

Aunque el estudio de las ecuaciones diferenciales comenzó en el siglo XVII, no fue hasta el XIX que los matemáticos se dieron cuenta que un número relativamente pequeño de ecuaciones diferenciales podía resolverse con medios elementales. Los trabajos de Cauchy, Liouville y otros, mostraron la importancia de establecer teoremas generales para garantizar la existencia de soluciones para ciertas clases especiales de ecuaciones diferenciales. El capítulo 6 se dedicó al empleo de un teorema de existencia y unicidad en el estudio de las ecuaciones diferenciales lineales. En este capítulo nos interesamos en una demostración de ese teorema y en temas relacionados con él.

La teoría de existencia para ecuaciones diferenciales de orden superior puede reducirse al caso de primer orden introduciendo *sistemas* de ecuaciones. Por ejemplo, la ecuación de segundo orden

$$y'' + 2ty' - y = e^t \tag{7.1}$$

puede transformarse en un sistema de dos ecuaciones de primer orden introduciendo dos funciones incógnitas y_1 e y_2, siendo

$$y_1 = y\,, \qquad y_2 = y_1'\,.$$

Tenemos entonces $y_2' = y_1'' = y''$, así que (7.1) puede escribirse como un sistema de dos ecuaciones de primer orden:

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_1 - 2ty_2 + e^t\,. \end{aligned} \tag{7.2}$$

No podemos resolver las ecuaciones separadamente con los métodos del capítulo 6 porque cada una de ellas contiene dos funciones incógnitas.

En este capítulo consideramos sistemas que constan de n ecuaciones diferenciales lineales de primer orden con n funciones incógnitas $y_1, \ldots, y_n$. Esos sistemas tienen la forma

$$
\begin{aligned}
y_1' &= p_{11}(t)y_1 + p_{12}(t)y_2 + \cdots + p_{1n}(t)y_n + q_1(t) \\
&\;\;\vdots \\
y_n' &= p_{n1}(t)y_1 + p_{n2}(t)y_2 + \cdots + p_{nn}(t)y_n + q_n(t).
\end{aligned} \tag{7.3}
$$

Las funciones p_{ik} y q_i que aparecen en (7.3) se consideran funciones dadas definidas en un cierto intervalo J. Las funciones $y_1, \ldots, y_n$ son las funciones incógnitas que hay que determinar. Los sistemas de este tipo se llaman *sistemas lineales de primer orden.* En general cada ecuación del sistema contiene más de una función incógnita por lo cual las ecuaciones no pueden resolverse separadamente.

Una ecuación diferencial lineal de orden n siempre puede transformarse en un sistema. Supongamos que la ecuación de orden n dada es

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = R(t), \tag{7.4}$$

donde los coeficientes a_i son funciones dadas. Para transformarla en un sistema escribimos $y_1 = y$ e introducimos una nueva función incógnita para cada una de las sucesivas derivadas de y. Esto es, ponemos

$$y_1 = y, \qquad y_2 = y_1', \qquad y_3 = y_2', \ldots, y_n = y_{n-1}',$$

y escribimos (7.4) como sistema

$$
\begin{aligned}
y_1' &= y_2 \\
y_2' &= y_3 \\
&\;\;\vdots \\
y_{n-1}' &= y_n \\
y_n' &= -a_n y_1 - a_{n-1} y_2 - \cdots - a_1 y_n + R(t).
\end{aligned} \tag{7.5}
$$

La discusión de los sistemas puede simplificarse considerablemente mediante la notación matricial y vectorial. Consideremos el sistema general (7.3) e intro-

duzcamos las funciones vectoriales $Y = (y_1, \ldots, y_n)$, $Q = (q_1, \ldots, q_n)$, y una función matricial $P = [p_{ij}]$, definidas por las ecuaciones

$$Y(t) = (y_1(t), \ldots, y_n(t)), \qquad Q(t) = (q_1(t), \ldots, q_n(t)), \qquad P(t) = [p_{ji}(t)]$$

para cada t de J. Consideremos los vectores como matrices columna $n \times 1$ y escribamos el sistema (7.3) en la forma más sencilla

$$Y' = P(t)Y + Q(t). \tag{7.6}$$

Por ejemplo, en el sistema (7.2) tenemos

$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \qquad P(t) = \begin{bmatrix} 0 & 1 \\ 1 & -2t \end{bmatrix}, \qquad Q(t) = \begin{bmatrix} 0 \\ e^t \end{bmatrix}.$$

En el sistema (7.5) tenemos

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}, \qquad P(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}, \qquad Q(t) = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ R(t) \end{bmatrix}.$$

Un problema de valores iniciales para el sistema (7.6) es el de encontrar una función vectorial Y que satisfaga (7.6) y también una condición inicial de la forma $Y(a) = B$, donde $a \in J$ y $B = (b_1, \ldots, b_n)$ es un vector n-dimensional dado.

En el caso $n = 1$ (el caso escalar) sabemos según el teorema 6.1 que, si P y Q son continuas en J, todas las soluciones de (7.6) vienen dadas por la fórmula

$$Y(x) = e^{A(x)}Y(a) + e^{A(x)} \int_a^x e^{-A(t)}Q(t)\,dt, \tag{7.7}$$

donde $A(x) = \int_a^x P(t)\,dt$, y a es un punto cualquiera de J. Demostraremos que esta fórmula puede generalizarse en forma adecuada para sistemas, esto es, cuando $P(t)$ es una función matricial $n \times n$ y $Q(t)$ una función vectorial de dimensión n. Para ello tenemos que asignar un significado a las integrales de matrices y a las

exponenciales de matrices. Por consiguiente, vamos a comentar brevemente el cálculo con funciones matriciales.

7.2 Cálculo con funciones matriciales

La generalización de los conceptos de integral y derivada para funciones matriciales es completamente correcta. Si $P(t) = [p_{ij}(t)]$, definimos la integral $\int_a^b P(t)\,dt$ por

$$\int_a^b P(t)\,dt = \left[\int_a^b p_{ij}(t)\,dt\right].$$

Esto es, la integral de la matriz $P(t)$ es la matriz obtenida integrando cada elemento de $P(t)$, suponiendo como es natural, que cada elemento sea integrable en $[a, b]$. El lector puede comprobar que la linealidad para las integrales se generaliza a las funciones matriciales.

La continuidad y derivabilidad de funciones matriciales se definen también en función de los elementos. Decimos que una función matricial $P = [p_{ij}]$ es continua en t si cada elemento p_{ij} es continuo en t. La derivada P' se define derivando cada elemento

$$P'(t) = [p'_{ij}(t)],$$

siempre que existan todas las derivadas $p'_{ij}(t)$. Se comprueban fácilmente las reglas de derivación básicas para sumas y productos. Por ejemplo, si P y Q son funciones matriciales derivables, tenemos

$$(P + Q)' = P' + Q'$$

si P y Q son del mismo tamaño, y si el producto PQ está definido, tenemos también

$$(PQ)' = PQ' + P'Q$$

La regla de la cadena también es válida. Esto es, si $F(t) = P[g(t)]$, donde P es una función matricial derivable y g es una función escalar derivable, entonces $F'(t) = g'(t)P'[g(t)]$. El teorema de la derivada nula, y el primer y segundo teoremas fundamentales del cálculo también son válidos para funciones matriciales. Las demostraciones de esas propiedades se proponen en el próximo conjunto de ejercicios.

La definición de la exponencial de una matriz no es tan sencilla y exige más preparación. Tal concepto se expone en la sección que sigue.

7.3 Series de matrices. Normas de matrices

Sea $A = [a_{ij}]$ una matriz $n \times n$ de elementos reales o complejos. Queremos definir la exponencial e^A de manera que posea alguna de las propiedades fundamentales de la exponencial ordinaria de valores reales o complejos. En particular, exigiremos la ley de exponentes en la forma

(7.8) $$e^{tA}e^{sA} = e^{(t+s)A} \quad \text{para todo par } s \text{ y } t \text{ reales,}$$

y la relación

(7.9) $$e^O = I,$$

donde O e I son las matrices, $n \times n$, cero e identidad respectivamente. Puede parecer natural definir e^A como la matriz $[e^{a_{ij}}]$. No obstante, eso no es aceptable porque no satisface ninguna de las propiedades (7.8) o (7.9). En lugar de ello, definiremos e^A por medio de un desarrollo en serie de potencias,

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

Sabemos que esta fórmula es válida si A es un número real o complejo, y demostraremos que se satisfacen las propiedades (7.8) y (7.9) si A es una matriz. Antes necesitamos conocer qué se entiende por serie convergente de matrices.

DEFINICIÓN DE SERIE CONVERGENTE DE MATRICES. *Dada una sucesión $\{C_k\}$ de matrices $m \times n$ cuyos elementos son números reales o complejos, designemos el elemento ij de C_k por $c_{ij}^{(k)}$. Si todas las mn series*

(7.10) $$\sum_{k=1}^{\infty} c_{ij}^{(k)} \qquad (i = 1, \ldots, m; j = 1, \ldots, n)$$

son convergentes, decimos entonces que la serie de matrices $\sum_{k=1}^{\infty} C_k$ es convergente, y su suma está definida como la matriz $m \times n$ cuyo elemento ij es la serie (7.10).

Un sencillo y útil criterio de convergencia de una serie de matrices puede darse en función de la *norma* de una matriz, generalización del valor absoluto de un número.

DEFINICIÓN DE NORMA DE UNA MATRIZ. *Si $A = [a_{ij}]$ es una matriz $m \times n$*

de elementos reales o complejos, la norma de A, designada por $\|A\|$, *se define como el número no negativo dado por la fórmula*

$$(7.11) \qquad \|A\| = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}| .$$

Es decir, la norma de A es la suma de los valores absolutos de todos sus elementos. Algunas veces se usan otras definiciones de la norma, pero hemos elegido ésta debido a la facilidad con la que podemos demostrar las propiedades siguientes.

TEOREMA 7.1. PROPIEDADES FUNDAMENTALES DE LAS NORMAS. *Para las matrices rectangulares A y B, y todos los escalares c reales o complejos tenemos*

$$\|A + B\| \leq \|A\| + \|B\|, \ \|AB\| \leq \|A\| \ \|B\|, \ \|cA\| = |c| \ \|A\| .$$

Demostración. Solamente probamos el resultado para $\|AB\|$, suponiendo que A es $m \times n$ y B es $n \times p$. Las demostraciones de los otros son sencillas y las dejamos como ejercicios.

Escribiendo $A = [a_{ik}], B = [b_{kj}],$ tenemos $AB = [\ \sum_{k=1}^{n} a_{ik} b_{kj}]$, así que de (7.11) obtenemos

$$\|AB\| = \sum_{i=1}^{m} \sum_{j=1}^{p} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right| \leq \sum_{i=1}^{m} \sum_{k=1}^{n} |a_{ik}| \sum_{j=1}^{p} |b_{kj}| \leq \sum_{i=1}^{m} \sum_{k=1}^{n} |a_{ik}| \ \|B\| = \|A\| \ \|B\| .$$

Obsérvese que en el caso especial $B = A$ la desigualdad para $\|AB\|$ se convierte en $\|A^2\| \leq \|A\|^2$. Por inducción también tenemos

$$\|A^k\| \leq \|A\|^k \quad \text{para} \quad k = 1, 2, 3, \ldots .$$

Estas desigualdades se usarán en la discusión de la exponencial de una matriz.

El teorema que sigue nos da una útil condición suficiente para la convergencia de una serie de matrices.

TEOREMA 7.2. CRITERIO DE CONVERGENCIA PARA SERIES DE MATRICES. *Si* $\{C_k\}$ *es una sucesión de matrices* $m \times n$ *tales que* $\sum_{k=1}^{\infty} \|C_k\|$ *converge, entonces la serie de matrices* $\sum_{k=1}^{\infty} C_k$ *también converge.*

Demostración. Designemos el elemento ij de C_k por $c_{ij}^{(k)}$. Puesto que

$|c_{ij}^{(k)}| \leq \|C_k\|$, la convergencia de $\sum_{k=1}^{\infty} \|C_k\|$ implica la convergencia absoluta de cada una de las series $\sum_{k=1}^{\infty} c_{ij}^{(k)}$. Luego cada una de las series $\sum_{k=1}^{\infty} c_{ij}^{(k)}$ es convergente, por lo que la serie de matrices $\sum_{k=1}^{\infty} C_k$ es convergente.

7.4 Ejercicios

1. Comprobar que la linealidad de integrales también es válida para las integrales de funciones matriciales.
2. Comprobar cada una de las siguientes reglas de derivación para funciones matriciales, suponiendo que P y Q son derivables. En a), P y Q deben ser del mismo tamaño de manera que $P + Q$ tenga sentido. En b) y d) no deben ser necesariamente del mismo tamaño con tal que los productos tengan sentido. En c) y d), Q se supone no singular.

   a) $(P + Q)' = P' + Q'$.
   b) $(PQ)' = PQ' + P'Q$.
   c) $(Q^{-1})' = -Q^{-1}Q'Q^{-1}$.
   d) $(PQ^{-1})' = -PQ^{-1}Q'Q^{-1} + P'Q^{-1}$.
3. a) Sea P una función matricial derivable. Demostrar que las derivadas de P^2 y P^3 vienen dadas por las fórmulas

$$(P^2)' = PP' + P'P, \qquad (P^3)' = P^2P' + PP'P + P'P^2.$$

   b) Enunciar una fórmula general para la derivada de P^k y demostrarla por inducción.
4. Sean P una función matricial derivable y g una función escalar derivable cuyo recorrido sea un subconjunto del dominio de P. Definir la función compuesta $F(t) = P[g(t)]$ y demostrar la regla de la cadena, $F'(t) = g'(t)P'[g(t)]$.
5. Demostrar el teorema de la derivada nula para funciones matriciales: *Si $P'(t) = O$ para todo t en un intervalo abierto (a, b), la función matricial P es constante en (a, b).*
6. Establecer y demostrar generalizaciones del primero y segundo teoremas fundamentales del cálculo para funciones matriciales.
7. Establecer y demostrar una fórmula de integración por partes en la que los integrandos sean funciones matriciales.
8. Demostrar las propiedades siguientes de las normas matriciales:

$$\|A + B\| \leq \|A\| + \|B\|, \qquad \|cA\| = |c|\,\|A\|.$$

9. Si una función matricial P es integrable en un intervalo $[a, b]$ demostrar que

$$\left\|\int_a^b P(t)\, dt\right\| \leq \int_a^b \|P(t)\|\, dt.$$

10. Sea D una matriz diagonal $n \times n$, $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Demostrar que la serie matricial $\sum_{k=0}^{\infty} D^k/k!$ converge y es también una matriz diagonal,

$$\sum_{k=0}^{\infty} \frac{D^k}{k!} = \text{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}).$$

(El término correspondiente a $k = 0$ se sobreentiende que es la matriz identidad I.)

11. Sea D una matriz diagonal $n \times n$, $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Si la serie matricial $\sum_{k=0}^{\infty} c_k D^k$ converge, demostrar que

$$\sum_{k=0}^{\infty} c_k D^k = \text{diag}\left(\sum_{k=0}^{\infty} c_k \lambda_1^k, \ldots, \sum_{k=0}^{\infty} c_k \lambda_n^k\right).$$

12. Supongamos que la serie matricial $\sum_{k=1}^{\infty} C_k$ converge, donde cada C_k es una matriz $n \times n$. Demostrar que la serie matricial $\sum_{k=1}^{\infty} (AC_kB)$ también converge y que su suma es la matriz

$$A\left(\sum_{k=1}^{\infty} C_k\right)B.$$

Aquí A y B son matrices tales que los productos AC_kB tienen sentido.

7.5 Exponencial de una matriz

Aplicando el teorema 7.2 es fácil demostrar que la serie matricial

$$\sum_{k=0}^{\infty} \frac{A^k}{k!} \tag{7.12}$$

converge para cualquier matriz cuadrada A con elementos reales o complejos. (Se sobreentiende que el término correspondiente a $k = 0$ es la matriz identidad I.) La norma de cada uno de los términos satisface la desigualdad

$$\left\| \frac{A^k}{k!} \right\| \leq \frac{\|A\|^k}{k!}.$$

Puesto que la serie $\sum a^k/k!$ converge para todo número real a, el teorema 7.2 implica que la serie (7.12) converge para toda matriz cuadrada A.

DEFINICIÓN DE EXPONENCIAL DE UNA MATRIZ. *Dada una matriz cualquiera A, $n \times n$, con elementos reales o complejos definimos la exponencial e^A como la matriz $n \times n$ dada por la serie convergente* (7.12). *Esto es,*

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Obsérvese que esta definición implica $e^O = I$, donde O es la matriz cero. Con la ayuda de las ecuaciones diferenciales se estudiarán otras propiedades de la exponencial.

7.6 Ecuación diferencial que se satisface por e^{tA}

Sean t un número real, A una matriz $n \times n$ y $E(t)$ la matriz $n \times n$ dada por

$$E(t) = e^{tA}.$$

Mantendremos A fijo y estudiaremos esa matriz como una función de t. Obtenemos primero una ecuación diferencial a la que satisfaga E.

TEOREMA 7.3. *Para todo real t la función matricial E definida por $E(t) = e^{tA}$ satisface la ecuación diferencial matricial*

$$E'(t) = E(t)A = AE(t).$$

Demostración. De la definición de la exponencial de una matriz tenemos

$$E(t) = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}.$$

Designemos con $c_{ij}^{(k)}$ el elemento ij de A^k. Entonces el elemento ij de $t^k A^k/k!$ es $t^k c_{ij}^{(k)}/k!$. Luego, de la definición de serie matricial, tenemos

$$(7.13) \qquad \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = \left[\sum_{k=0}^{\infty} \frac{t^k}{k!} c_{ij}^{(k)} \right].$$

Cada elemento del segundo miembro en (7.13) es una serie de potencias en t, convergente para todo t y viene dada por la serie derivada

$$\sum_{k=1}^{\infty} \frac{k t^{k-1}}{k!} c_{ij}^{(k)} = \sum_{k=0}^{\infty} \frac{t^k}{k!} c_{ij}^{(k+1)}.$$

Esto demuestra que existe la derivada $E'(t)$ y viene dada por la serie matricial

$$E'(t) = \sum_{k=0}^{\infty} \frac{t^k A^{k+1}}{k!} = \left(\sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \right) A = E(t)A.$$

En la última ecuación utilizamos la propiedad de que $A^{k+1} = A^k A$. Puesto que A es permutable con A^k podríamos también escribir $A^{k+1} = AA^k$ para obtener la relación $E'(t) = AE(t)$. Esto completa la demostración.

Observación: La demostración anterior pone también en evidencia que A es permutable con e^{tA}.

7.7 Teorema de unicidad para la ecuación diferencial matricial $F'(t) = AF(t)$

En esta sección demostramos un teorema de unicidad que caracteriza todas las soluciones de la ecuación diferencial matricial $F'(t) = AF(t)$. La demostración utiliza el teorema siguiente.

TEOREMA 7.4. NO SINGULARIDAD DE e^{tA}. *Cualesquiera que sean la matriz A, $n \times n$, y el escalar t tenemos*

$$e^{tA}e^{-tA} = I. \tag{7.14}$$

Luego e^{tA} es no singular, y su inversa es e^{-tA}.

Demostración. Sea F la función matricial definida para todo número real t mediante la ecuación

$$F(t) = e^{tA}e^{-tA}.$$

Demostraremos que $F(t)$ es la matriz identidad I haciendo ver que la derivada $F'(t)$ es la matriz cero. Derivando F como un producto, y recordando el resultado del teorema 7.3, encontramos

$$\begin{aligned} F'(t) &= e^{tA}(e^{-tA})' + (e^{tA})'e^{-tA} = e^{tA}(-Ae^{-tA}) + Ae^{tA}e^{-tA} \\ &= -Ae^{tA}e^{-tA} + Ae^{tA}e^{-tA} = O, \end{aligned}$$

puesto que A es permutable con e^{tA}. Por consiguiente, según el teorema de la derivada nula, F es una matriz constante. Pero $F(0) = e^{0A}e^{0A} = I$, por lo cual $F(t) = I$ para todo t. Esto demuestra (7.14).

TEOREMA 7.5. TEOREMA DE UNICIDAD. *Sean A y B dos matrices constantes $n \times n$ dadas. La única función matricial F, $n \times n$, que satisface el problema de valor inicial*

$$F'(t) = AF(t), \qquad F(0) = B$$

para $-\infty < t < +\infty$ es

$$F(t) = e^{tA}B. \tag{7.15}$$

Demostración. Observemos primero que $e^{tA}B$ es una solución. Sea ahora F una solución cualquiera y consideremos la función matricial

$$G(t) = e^{-tA}F(t).$$

Derivando este producto obtenemos

$$G'(t) = e^{-tA}F'(t) - Ae^{-tA}F(t) = e^{-tA}AF(t) - e^{-tA}AF(t) = O.$$

Por consiguiente $G(t)$ es una matriz constante,

$$G(t) = G(0) = F(0) = B.$$

Es decir, $e^{-tA}F(t) = B$. Multiplicando por e^{tA} y aplicando (7.14) obtenemos (7.15).

Observación: El mismo tipo de demostración hace ver que $F(t) = Be^{tA}$ es la única solución del problema de valor inicial

$$F'(t) = F(t)A, \qquad F(0) = B.$$

7.8 Ley de exponentes para exponenciales de matrices

La ley de exponentes $e^Ae^B = e^{A+B}$ no siempre es válida para exponenciales de matrices. En el ejercicio 13 de la sección 7.12 se da un contraejemplo. No obstante, no es difícil demostrar que la fórmula es válida para matrices permutables.

TEOREMA 7.6. *Sean A y B dos matrices $n\times n$ permutables, $AB = BA$. Tenemos*

$$e^{A+B} = e^Ae^B. \tag{7.16}$$

Demostración. De la ecuación $AB = BA$ deducimos que

$$A^2B = A(BA) = (AB)A = (BA)A = BA^2,$$

de modo que B es permutable con A^2. Por inducción, B es permutable con cualquier potencia de A. Escribiendo e^{tA} en forma de serie de potencias encontramos que B también es permutable con e^{tA} para todo t real.

Sea ahora F la función matricial definida por la ecuación

$$F(t) = e^{t(A+B)} - e^{tA}e^{tB}.$$

Derivando $F(t)$ y teniendo en cuenta que B es permutable con e^{tA} encontramos

$$\begin{aligned}F'(t) &= (A + B)e^{t(A+B)} - Ae^{tA}e^{tB} - e^{tA}Be^{tB}\\ &= (A + B)e^{t(A+B)} - (A + B)e^{tA}e^{tB} = (A + B)F(t).\end{aligned}$$

Según el teorema de unicidad tenemos

$$F(t) = e^{t(A+B)}F(0).$$

Pero $F(0) = O$, así que $F(t) = O$ para todo t. Luego

$$e^{t(A+B)} = e^{tA}e^{tB}.$$

Cuando $t = 1$ obtenemos (7.16).

EJEMPLO. Las matrices sA y tA son permutables para todos los escalares s y t. Luego tenemos

$$e^{sA}e^{tA} = e^{(s+t)A}.$$

7.9 Teoremas de existencia y unicidad para sistemas lineales homogéneos con coeficientes constantes

La ecuación diferencial vectorial $Y'(t) = AY(t)$, donde A es una matriz constante $n \times n$ e Y es una función vectorial n-dimensional (considerada como una matriz columna $n \times 1$) se denomina *sistema lineal homogéneo con coeficientes constantes.* Utilizaremos la exponencial de una matriz para dar una fórmula explícita de la solución de un tal sistema.

TEOREMA 7.7. *Sean A una matriz constante $n \times n$ dada y B un vector n-dimensional. El problema de valor inicial*

$$Y'(t) = AY(t), \qquad Y(0) = B, \tag{7.17}$$

tiene solución única en el intervalo $-\infty < t < +\infty$. Esta solución viene dada por la fórmula

$$Y(t) = e^{tA}B. \tag{7.18}$$

Más general, la única solución del problema de valor inicial

$$Y'(t) = AY(t), \qquad Y(a) = B,$$

es $Y(t) = e^{(t-a)A}B$.

Demostración. La derivación de (7.18) nos da $Y'(t) = Ae^{tA}B = AY(t)$. Pues-

to que $Y(0) = B$, ésta es una solución del problema de valor inicial (7.17).

Para demostrar que es la única solución, razonamos como en la demostración del teorema 7.5. Sea $Z(t)$ otra función vectorial que satisfaga $Z'(t) = AZ(t)$ con $Z(0) = B$, y pongamos $G(t) = e^{-tA}Z(t)$. Comprobamos entonces con facilidad que $G'(t) = O$, así que $G(t) = G(0) = Z(0) = B$. Es decir, $e^{-tA}Z(t) = B$, de modo que $Z(t) = e^{tA}B = Y(t)$. Del mismo modo se puede discutir el caso más general con valor inicial $Y(a) = B$.

7.10 El problema de calcular e^{tA}

Aunque el teorema 7.7 da una fórmula explícita para la solución de un sistema homogéneo con coeficientes constantes, queda todavía el problema de calcular efectivamente la exponencial e^{tA}. Si se tratara de calcular e^{tA} directamente a partir de la definición de serie tendríamos que calcular todas las potencias A^k para $k = 0, 1, 2, \ldots$, y calcular luego la suma de cada serie $\sum_{k=0}^{\infty} t^k c_{ij}^{(k)}/k!$, donde $c_{ij}^{(k)}$ es el elemento ij de A^k. En general ése es un trabajo desesperante salvo si A es una matriz cuyas potencias pueden calcularse fácilmente. Por ejemplo, si A es una matriz diagonal, sea

$$A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

entonces toda potencia de A también es una matriz diagonal, pues

$$A^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k).$$

Por consiguiente en este caso e^{tA} es una matriz diagonal dada por

$$e^{tA} = \operatorname{diag}\left(\sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_1^k, \ldots, \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_n^k\right) = \operatorname{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n}).$$

Otro caso fácil de manejar es cuando A es una matriz que puede diagonalizarse. Por ejemplo, si existe una no singular C tal que $C^{-1}AC$ es una matriz diagonal, sea $C^{-1}AC = D$, entonces tenemos $A = CDC^{-1}$, a partir de la cual encontramos

$$A^2 = (CDC^{-1})(CDC^{-1}) = CD^2C^{-1},$$

y, más general,

$$A^k = CD^kC^{-1}.$$

Por tanto en este caso tenemos

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = \sum_{k=0}^{\infty} \frac{t^k}{k!} CD^kC^{-1} = C\left(\sum_{k=0}^{\infty} \frac{t^k D^k}{k!}\right)C^{-1} = Ce^{tD}C^{-1}.$$

Aquí la dificultad está en determinar C y su inversa. Una vez ésas son conocidas, e^{tA} se calcula con facilidad. Naturalmente, no toda matriz puede diagonalizarse de modo que la utilidad de las observaciones precedentes es limitada.

EJEMPLO 1. Calcular e^{tA} para la matriz 2×2, $A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$.

Solución. Esta matriz tiene autovalores distintos $\lambda_1 = 6$, $\lambda_2 = 1$, así que existe una matriz no singular $C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tal que $C^{-1}AC = D$, donde $D = \operatorname{diag}(\lambda_1, \lambda_2) = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$. Para determinar C podemos escribir $AC = CD$, o

$$\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}.$$

Multiplicando las matrices, encontramos que esa ecuación se satisface cualesquiera que sean los escalares a, b, c, d tales que $a = 4c$, $b = -d$. Tomando $c = d = 1$ elegimos

$$C = \begin{bmatrix} 4 & -1 \\ 1 & 1 \end{bmatrix}, \qquad C^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 1 \\ -1 & 4 \end{bmatrix}.$$

Por lo tanto

$$e^{tA} = Ce^{tD}C^{-1} = \frac{1}{5}\begin{bmatrix} 4 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} e^{6t} & 0 \\ 0 & e^t \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 4 \end{bmatrix}$$

$$= \frac{1}{5}\begin{bmatrix} 4 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} e^{6t} & e^{6t} \\ -e^t & 4e^t \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4e^{6t} + e^t & 4e^{6t} - 4e^t \\ e^{6t} - e^t & e^{6t} + 4e^t \end{bmatrix}.$$

EJEMPLO 2. Resolver el sistema lineal

$$y_1' = 5y_1 + 4y_2$$
$$y_2' = y_1 + 2y_2$$

sujeto a las condiciones iniciales $y_1(0) = 2$, $y_2(0) = 3$.

Solución. En forma matricial el sistema se puede escribir así:

$$Y'(t) = AY(t), \qquad Y(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \qquad \text{donde} \quad A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}.$$

Según el teorema 7.7 la solución es $Y(t) = e^{tA}Y(0)$. Usando la matriz e^{tA} calculada en el ejemplo 1 encontramos

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4e^{6t} + e^t & 4e^{6t} - 4e^t \\ e^{6t} - e^t & e^{6t} + 4e^t \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

de la que obtenemos

$$y_1 = 4e^{6t} - 2e^t, \qquad y_2 = e^{6t} + 2e^t.$$

Se conocen muchos métodos para calcular e^{tA} cuando A no puede diagonalizarse. La mayoría de ellos son más bien complicados y exigen transformaciones matriciales previas, cuya naturaleza depende de las multiplicidades de los autovalores de A. Más adelante discutiremos un método práctico y seguro para calcular e^{tA} que se puede utilizar tanto si A puede diagonalizarse como si no. Es válido para *todas* las matrices A y no exige transformaciones previas de ningún tipo. Tal método fue desarrollado por E. J. Putzer en un artículo del *American Mathematical Monthly*, Vol. 73 (1966), págs. 2-7. Se basa en un famoso teorema atribuido a Arthur Cayley (1821-1895) y William Rowan Hamilton (1805-1865) que establece que toda matriz cuadrada satisface su ecuación característica. Demostraremos primero el teorema de Cayley-Hamilton y luego lo aplicaremos para obtener la fórmula de Putzer para calcular e^{tA}.

7.11 Teorema de Cayley-Hamilton

TEOREMA 7.8. TEOREMA DE CAYLEY-HAMILTON. *Sean A una matriz $n \times n$ y*

$$f(\lambda) = \det(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 \tag{7.19}$$

su polinomio característico. Entonces $f(A) = O$. Es decir, A satisface la ecuación

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I = O. \tag{7.20}$$

Demostración. La demostración se basa en el teorema 3.12 que establece que para toda matriz cuadrada A se tiene

$$A(\operatorname{cof} A)^t = (\det A)I. \tag{7.21}$$

Apliquemos esta fórmula reemplazando A por $\lambda I - A$. Puesto que $\det(\lambda I - A) = f(\lambda)$, la ecuación (7.21) se convierte en

$$(\lambda I - A)\{\operatorname{cof}(\lambda I - A)\}^t = f(\lambda)I. \tag{7.22}$$

Esta ecuación es válida para todo real λ. La idea de la demostración consiste en hacer ver que también es válida cuando λ se reemplaza por A.

Los elementos de la matriz $\operatorname{cof}(\lambda I - A)$ son los cofactores de $\lambda I - A$. Excepto para un factor ± 1, cada uno de tales cofactores es el determinante de una menor de $\lambda I - A$ de orden $n - 1$. Por lo tanto cada elemento de $\operatorname{cof}(\lambda I - A)$, y por tanto de $\{\operatorname{cof}(\lambda I - A)\}^t$, es un polinomio en λ de grado $\leq n - 1$. Por consiguiente

$$\{\operatorname{cof}(\lambda I - A)\}^t = \sum_{k=0}^{n-1} \lambda^k B_k,$$

donde cada coeficiente B_k es una matriz $n \times n$ con elementos escalares. Aplicando esto en (7.22) obtenemos la relación

$$(\lambda I - A)\sum_{k=0}^{n-1} \lambda^k B_k = f(\lambda)I \tag{7.23}$$

que se puede poner en la forma

$$\lambda^n B_{n-1} + \sum_{k=1}^{n-1} \lambda^k (B_{k-1} - AB_k) - AB_0 = \lambda^n I + \sum_{k=1}^{n-1} \lambda^k c_k I + c_0 I. \tag{7.24}$$

Igualando ahora los coeficientes de las potencias semejantes de λ en (7.24) obtenemos las ecuaciones

$$\begin{aligned} B_{n-1} &= I \\ B_{n-2} - AB_{n-1} &= c_{n-1}I \\ &\;\;\vdots \\ B_0 - AB_1 &= c_1 I \\ -AB_0 &= c_0 I. \end{aligned} \tag{7.25}$$

La igualación de los coeficientes es posible porque (7.24) es equivalente a n^2 ecuaciones escalares, en cada una de las cuales podemos igualar los coeficientes de potencias semejantes de λ. Multipliquemos a continuación las ecuaciones (7.25)

sucesivamente por $A^n, A^{n-1}, \ldots, A, I$ y sumemos los resultados. Los términos del primer miembro desaparecen y obtenemos

$$O = A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I.$$

lo cual demuestra el teorema de Cayley-Hamilton.

Observación: Hamilton demostró el teorema en 1853 para un tipo especial de matrices. Pocos años más tarde, Cayley anunció que el teorema era válido para todas las matrices, pero no dio la demostración

EJEMPLO. La matriz $A = \begin{bmatrix} 5 & 4 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 2 \end{bmatrix}$ tiene como polinomio característico

$$f(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 6) = \lambda^3 - 9\lambda^2 + 20\lambda - 12.$$

El teorema de Cayley-Hamilton establece que A satisface la ecuación

$$A^3 - 9A^2 + 20A - 12I = O. \tag{7.26}$$

Esta ecuación puede usarse para expresar A^3 y todas las potencias superiores de A en función de I, A y A^2. Por ejemplo, tenemos

$$\begin{aligned} A^3 &= 9A^2 - 20A + 12I, \\ A^4 &= 9A^3 - 20A^2 + 12A = 9(9A^2 - 20A + 12I) - 20A^2 + 12A \\ &= 61A^2 - 168A + 108I. \end{aligned}$$

También se puede emplear para expresar A^{-1} como un polinomio en A. De (7.26) resulta también $A(A^2 - 9A + 20I) = 12I$, y obtenemos

$$A^{-1} = \tfrac{1}{12}(A^2 - 9A + 20I).$$

7.12 Ejercicios

En cada uno de los ejercicios del 1 al 4, a) expresar A^{-1}, A^2 y todas las potencias supe-

riores de A como combinaciones lineales de I y A. (El teorema de Cayley-Hamilton puede servir de ayuda). b) Calcular e^{tA}.

1. $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. 2. $A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$. 3. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. 4. $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

5. a) Si $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, demostrar que $e^{tA} = \begin{bmatrix} \cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t \end{bmatrix}$.

   b) Hallar la fórmula correspondiente para e^{tA} cuando $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, a, b reales.

6. Si $F(t) = \begin{bmatrix} t & t-1 \\ 0 & 1 \end{bmatrix}$, demostrar que $e^{F(t)} = eF(e^{t-1})$.

7. Si $A(t)$ es una función escalar de t, la derivada de $e^{A(t)}$ es $e^{A(t)}A'(t)$. Calcular la derivada de $e^{A(t)}$ cuando $A(t) = \begin{bmatrix} 1 & t \\ 0 & 0 \end{bmatrix}$ y demostrar que el resultado no es igual a ninguno de los dos productos $e^{A(t)}A'(t)$ o $A'(t)e^{A(t)}$.

En cada uno de los ejercicios 8, 9, 10, a) calcular A^n, y expresar A^3 en función de I, A, A^2. b) Calcular e^{tA}.

8. $A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. 9. $A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. 10. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

11. Si $A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, expresar e^{tA} como combinación lineal de I, A, A^2.

12. Si $A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$ demostrar que $e^{A} = \begin{bmatrix} x^2 & xy & y^2 \\ 2xy & x^2 + y^2 & 2xy \\ y^2 & xy & x^2 \end{bmatrix}$, donde $x = \cosh 1$ e $y = \operatorname{senh} 1$.

13. Este ejemplo demuestra que la ecuación $e^{A+B} = e^{A}e^{B}$ no siempre es válida para las exponenciales de matrices. Calcular cada una de las matrices $e^{A}e^{B}$, $e^{B}e^{A}$, e^{A+B} cuando

$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$, y observar que los tres resultados son distintos.

7.13 Método de Putzer para calcular e^{tA}

El teorema de Cayley-Hamilton demuestra que la potencia n-ésima de cualquier matriz A, $n \times n$, puede expresarse como combinación lineal de las potencias inferiores $I, A, A^2, \ldots, A^{n-1}$. De aquí resulta que cada una de las potencias superiores $A^{n+1}, A^{n+2}, \ldots$, también puede expresarse como combinación lineal de $I, A, A^2, \ldots, A^{n-1}$. Por lo tanto, en la serie que define e^{tA}, cada término $t^k A^k / k!$ con $k \geq n$ es una combinación lineal de $t^k I, t^k A, t^k A^2, \ldots, t^k A^{n-1}$. Luego cabe pensar que e^{tA} puede expresarse como un polinomio en A de la forma

(7.27)
$$e^{tA} = \sum_{k=0}^{n-1} q_k(t) A^k,$$

donde los coeficientes escalares $q_k(t)$ dependen de t. Putzer desarrolló dos métodos útiles para expresar e^{tA} como un polinomio en A. El teorema que sigue describe el más sencillo de los dos métodos.

TEOREMA 7.9. *Sean $\lambda_1, \ldots, \lambda_n$ los autovalores de una matriz A, $n \times n$, y definamos una sucesión de polinomios en A como sigue:*

(7.28)
$$P_0(A) = I, \qquad P_k(A) = \prod_{m=1}^{k} (A - \lambda_m I), \quad \textit{para} \quad k = 1, 2, \ldots, n.$$

Tenemos entonces

(7.29)
$$e^{tA} = \sum_{k=0}^{n-1} r_{k+1}(t) P_k(A),$$

donde los coeficientes escalares $r_1(t), \ldots, r_n(t)$ se determinan por recurrencia a partir del sistema de ecuaciones diferenciales lineales.

(7.30)
$$\begin{aligned} r_1'(t) &= \lambda_1 r_1(t), \qquad r_1(0) = 1, \\ r_{k+1}'(t) &= \lambda_{k+1} r_{k+1}(t) + r_k(t), \qquad r_{k+1}(0) = 0, \qquad (k = 1, 2, \ldots, n-1). \end{aligned}$$

Observación: La ecuación (7.29) no expresa e^{tA} directamente en potencias de A como indica (7.27), sino como una combinación lineal de los polinomios $P_0(A), P_1(A), \ldots, P_{n-1}(A)$. Estos polinomios se calculan fácilmente en cuanto se han determinado los autovalores de A. Asimismo son de fácil cálculo los multiplicadores $r_1(t), \ldots, r_n(t)$ de (7.30). Aunque esto exige resolver un sistema de ecuacio-

nes diferenciales lineales, el sistema que nos ocupa tiene una matriz triangular y las soluciones pueden determinarse en forma sucesiva.

Demostración. Sean $r_1(t), \ldots, r_n(t)$ las funciones escalares determinadas por (7.30) y definamos una función matricial F por inducción

(7.31) $$F(t) = \sum_{k=0}^{n-1} r_{k+1}(t)P_k(A).$$

Obsérvese que $F(0) = r_1(0)P_0(A) = I$. Demostraremos que $F(t) = e^{tA}$ probando que F satisface la misma ecuación diferencial que e^{tA}, es decir, $F'(t) = AF(t)$.

Derivando (7.31) y usando las fórmulas de recurrencia (7.30) obtenemos

$$F'(t) = \sum_{k=0}^{n-1} r'_{k+1}(t)P_k(A) = \sum_{k=0}^{n-1} \{r_k(t) + \lambda_{k+1}r_{k+1}(t)\}P_k(A),$$

donde $r_0(t)$ es por definición 0. Volvemos a escribir esa fórmula en la forma

$$F'(t) = \sum_{k=0}^{n-2} r_{k+1}(t)P_{k+1}(A) + \sum_{k=0}^{n-1} \lambda_{k+1}r_{k+1}(t)P_k(A),$$

restamos luego $\lambda_n F(t) = \sum_{k=0}^{n-1} \lambda_n r_{k+1}(t)P_k(A)$ para obtener la relación

(7.32) $$F'(t) - \lambda_n F(t) = \sum_{k=0}^{n-2} r_{k+1}(t)\{P_{k+1}(A) + (\lambda_{k+1} - \lambda_n)P_k(A)\}.$$

Pero de (7.28) vemos que $P_{k+1}(A) = (A - \lambda_{k+1}I)P_k(A)$, por lo que

$$P_{k+1}(A) + (\lambda_{k+1} - \lambda_n)P_k(A) = (A - \lambda_{k+1}I)P_k(A) + (\lambda_{k+1} - \lambda_n)P_k(A)$$
$$= (A - \lambda_n I)P_k(A).$$

Por consiguiente la ecuación (7.32) se convierte en

$$F'(t) - \lambda_n F(t) = (A - \lambda_n I)\sum_{k=0}^{n-2} r_{k+1}(t)P_k(A) = (A - \lambda_n I)\{F(t) - r_n(t)P_{n-1}(A)\}$$
$$= (A - \lambda_n I)F(t) - r_n(t)P_n(A).$$

El teorema de Cayley-Hamilton implica que $P_n(A) = O$, así que la última ecuación se convierte en

$$F'(t) - \lambda_n F(t) = (A - \lambda_n I)F(t) = AF(t) - \lambda_n F(t),$$

de la que resulta $F'(t) = AF(t)$. Puesto que $F(0) = I$, el teorema de unicidad (7.7) demuestra que $F(t) = e^{tA}$.

EJEMPLO 1. Expresar e^{tA} como combinación lineal de I y de A si A es una matriz 2×2 con sus autovalores iguales a λ.

Solución. Poniendo $\lambda_1 = \lambda_2 = \lambda$, tenemos que resolver el sistema de ecuaciones diferenciales

$$r_1'(t) = \lambda r_1(t), \qquad r_1(0) = 1,$$

$$r_2'(t) = \lambda r_2(t) + r_1(t), \qquad r_2(0) = 0.$$

Resolviendo esas ecuaciones de primer orden en forma sucesiva encontramos

$$r_1(t) = e^{\lambda t}, \qquad r_2(t) = te^{\lambda t}.$$

Puesto que $P_0(A) = I$ y $P_1(A) = A - \lambda I$, la fórmula pedida para e^{tA} es

$$e^{tA} = e^{\lambda t}I + te^{\lambda t}(A - \lambda I) = e^{\lambda t}(1 - \lambda t)I + te^{\lambda t}A. \tag{7.33}$$

EJEMPLO 2. Resolver el ejemplo 1 si los autovalores de A son λ y μ, tales que $\lambda \neq \mu$.

Solución. En este caso el sistema de ecuaciones diferenciales es

$$r_1'(t) = \lambda r_1(t), \qquad r_1(0) = 1,$$
$$r_2'(t) = \mu r_2(t) + r_1(t), \qquad r_2(0) = 0.$$

Sus soluciones vienen dadas por

$$r_1(t) = e^{\lambda t}, \qquad r_2(t) = \frac{e^{\lambda t} - e^{\mu t}}{\lambda - \mu}.$$

Puesto que $P_0(A) = I$ y $P_1(A) = A - \lambda I$ la fórmula pedida para e^{tA} es

$$e^{tA} = e^{\lambda t}I + \frac{e^{\lambda t} - e^{\mu t}}{\lambda - \mu}(A - \lambda I) = \frac{\lambda e^{\mu t} - \mu e^{\lambda t}}{\lambda - \mu}I + \frac{e^{\lambda t} - e^{\mu t}}{\lambda - \mu}A. \tag{7.34}$$

Si los autovalores λ y μ son números complejos, las exponenciales $e^{\lambda t}$ y $e^{\mu t}$ también lo serán. Pero si λ y μ son complejos conjugados las expresiones que mul-

tiplican I y A en (7.34) serán reales. Por ejemplo, supongamos

$$\lambda = \alpha + i\beta, \qquad \mu = \alpha - i\beta, \qquad \beta \neq 0.$$

Entonces $\lambda - \mu = 2i\beta$ así que la ecuación (7.34) se transforma en

$$\begin{aligned} e^{tA} &= e^{(\alpha+i\beta)t}I + \frac{e^{(\alpha+i\beta)t} - e^{(\alpha-i\beta)t}}{2i\beta}[A - (\alpha + i\beta)I] \\ &= e^{\alpha t}\left\{e^{i\beta t}I + \frac{e^{i\beta t} - e^{-i\beta t}}{2i\beta}(A - \alpha I - i\beta I)\right\} \\ &= e^{\alpha t}\left\{(\cos\beta t + i\,\mathrm{sen}\,\beta t)I + \frac{\mathrm{sen}\,\beta t}{\beta}(A - \alpha I - i\beta I)\right\}. \end{aligned}$$

Los términos que contienen i desaparecen y queda

$$e^{tA} = \frac{e^{\alpha t}}{\beta}\{(\beta\cos\beta t - \alpha\,\mathrm{sen}\,\beta t)I + \mathrm{sen}\,\beta t\,A\}. \tag{7.35}$$

7.14 Otros métodos para calcular e^{tA} en casos especiales

El método de Putzer para expresar e^{tA} como un polinomio en A es completamente general debido a que es válido para todas las matrices cuadradas A. Un método general no siempre es el más sencillo en ciertos casos particulares. En esta sección damos métodos más sencillos para calcular e^{tA} en tres casos especiales: a) Cuando todos los autovalores de A son iguales, b) cuando todos los autovalores de A son distintos, y c) cuando A tiene dos autovalores distintos y uno de ellos precisamente de multiplicidad 1.

TEOREMA 7.10. *Si A es una matriz $n\times n$ con todos sus autovalores iguales a λ, entonces tenemos*

$$e^{tA} = e^{\lambda t}\sum_{k=0}^{n-1}\frac{t^k}{k!}(A - \lambda I)^k. \tag{7.36}$$

Demostración. Puesto que las matrices λtI y $t(A - \lambda I)$ son permutables tenemos

$$e^{tA} = e^{\lambda tI}e^{t(A-\lambda I)} = (e^{\lambda t}I)\sum_{k=0}^{\infty}\frac{t^k}{k!}(A - \lambda I)^k.$$

El teorema de Cayley-Hamilton implica que $(A - \lambda I)^k = O$ para todo $k \geq n$, así que el teorema queda demostrado.

TEOREMA 7.11. *Si A es una matriz $n \times n$ con n autovalores distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$, entonces tenemos*

$$e^{tA} = \sum_{k=1}^{n} e^{t\lambda_k} L_k(A),$$

donde $L_k(A)$ es un polinomio de grado $n - 1$ dado por la fórmula

$$L_k(A) = \prod_{\substack{j=1 \\ j \neq k}}^{n} \frac{A - \lambda_j I}{\lambda_k - \lambda_j} \quad para \quad k = 1, 2, \ldots, n.$$

Observación: **Los polinomios $L_k(A)$ se llaman** ***coeficientes de interpolación de Lagrange.***

Demostración. Definamos una función matricial F por la ecuación

$$F(t) = \sum_{k=1}^{n} e^{t\lambda_k} L_k(A) \tag{7.37}$$

y comprobemos que F satisface la ecuación diferencial $F'(t) = AF(t)$ y la condición inicial $F(0) = I$. De (7.37) vemos que

$$AF(t) - F'(t) = \sum_{k=1}^{n} e^{t\lambda_k}(A - \lambda_k I) L_k(A).$$

Según el teorema de Cayley-Hamilton tenemos $(A - \lambda_k I) L_k(A) = O$ para cada k, de modo que F satisface la ecuación diferencial $F'(t) = AF(t)$.

Para completar la demostración necesitamos probar que F satisface la condición inicial $F(0) = I$, que se convierte en

$$\sum_{k=1}^{n} L_k(A) = I. \tag{7.38}$$

En el ejercicio 16 de la sección 7.15 se esboza una demostración de (7.38).

El teorema que sigue trata el caso en el que A tiene dos autovalores distintos, uno de los cuales tiene precisamente multiplicidad 1.

TEOREMA 7.12. *Sea A una matriz $n \times n$ $(n \geq 3)$ con dos autovalores distintos λ y μ, tales que λ tiene multiplicidad $n - 1$ y μ tiene multiplicidad 1. Tenemos entonces*

$$e^{tA} = e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A-\lambda I)^k + \left\{\frac{e^{\mu t}}{(\mu-\lambda)^{n-1}} - \frac{e^{\lambda t}}{(\mu-\lambda)^{n-1}}\sum_{k=0}^{n-2}\frac{t^k}{k!}(\mu-\lambda)^k\right\}(A-\lambda I)^{n-1}.$$

Demostración. Como en el caso del teorema 7.10 empezamos escribiendo

$$e^{tA} = e^{\lambda t}\sum_{k=0}^{\infty}\frac{t^k}{k!}(A-\lambda I)^k = e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A-\lambda I)^k + e^{\lambda t}\sum_{k=n-1}^{\infty}\frac{t^k}{k!}(A-\lambda I)^k$$

$$= e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A-\lambda I)^k + e^{\lambda t}\sum_{r=0}^{\infty}\frac{t^{n-1+r}}{(n-1+r)!}(A-\lambda I)^{n-1+r}.$$

Calculemos ahora en forma finita la serie con índice r del segundo término del segundo miembro mediante el teorema de Cayley-Hamilton. Puesto que

$$A - \mu I = A - \lambda I - (\mu - \lambda)I$$

encontramos

$$(A-\lambda I)^{n-1}(A-\mu I) = (A-\lambda I)^n - (\mu-\lambda)(A-\lambda I)^{n-1}.$$

El primer miembro es O en virtud del teorema de Cayley-Hamilton así que

$$(A-\lambda I)^n = (\mu-\lambda)(A-\lambda I)^{n-1}.$$

Usando esta relación reiteradamente llegamos a

$$(A-\lambda I)^{n-1+r} = (\mu-\lambda)^r(A-\lambda I)^{n-1}.$$

Por consiguiente la serie con índice r se convierte en

$$\sum_{r=0}^{\infty}\frac{t^{n-1+r}}{(n-1+r)!}(\mu-\lambda)^r(A-\lambda I)^{n-1} = \frac{1}{(\mu-\lambda)^{n-1}}\sum_{k=n-1}^{\infty}\frac{t^k}{k!}(\mu-\lambda)^k(A-\lambda I)^{n-1}$$

$$= \frac{1}{(\mu-\lambda)^{n-1}}\left\{e^{t(\mu-\lambda)} - \sum_{k=0}^{n-2}\frac{t^k}{k!}(\mu-\lambda)^k\right\}(A-\lambda I)^{n-1}.$$

Esto completa la demostración.

La fórmula explícita del teorema 7.12 puede deducirse también aplicando el método de Putzer, pero los detalles son más complicados.

Las fórmulas explícitas de los teoremas 7.10, 7.11 y 7.12 sirven para todas las matrices de orden $n \leq 3$. Puesto que el caso 3×3 es frecuente en la práctica, resumimos a continuación las fórmulas para mayor facilidad en su utilización.

CASO I. Si una matriz A, 3×3, tiene autovalores $\lambda, \lambda, \lambda$, entonces

$$e^{tA} = e^{\lambda t}\{I + t(A - \lambda I) + \tfrac{1}{2}t^2(A - \lambda I)^2\}.$$

CASO II. Si una matriz A, 3×3, tiene autovalores λ, μ, ν, entonces

$$e^{tA} = e^{\lambda t}\frac{(A - \mu I)(A - \nu I)}{(\lambda - \mu)(\lambda - \nu)} + e^{\mu t}\frac{(A - \lambda I)(A - \nu I)}{(\mu - \lambda)(\mu - \nu)} + e^{\nu t}\frac{(A - \lambda I)(A - \mu I)}{(\nu - \lambda)(\nu - \mu)}.$$

CASO III. Si una matriz A, 3×3, tiene autovalores λ, λ, μ, siendo $\lambda \neq \mu$, entonces

$$e^{tA} = e^{\lambda t}\{I + t(A - \lambda I)\} + \frac{e^{\mu t} - e^{\lambda t}}{(\mu - \lambda)^2}(A - \lambda I)^2 - \frac{te^{\lambda t}}{\mu - \lambda}(A - \lambda I)^2.$$

EJEMPLO. Calcular e^{tA} cuando $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$.

Solución. Los autovalores de A son 1, 1, 2, así que la fórmula del caso III nos da

(7.39) $$e^{tA} = e^{t}\{I + t(A - I)\} + (e^{2t} - e^{t})(A - I)^2 - te^{t}(A - I)^2.$$

Haciendo operaciones y ordenando obtenemos

(7.40) $$e^{tA} = (-2te^{t} + e^{2t})I + \{(3t + 2)e^{t} - 2e^{2t}\}A - \{(t + 1)e^{t} - e^{2t}\}A^2.$$

Ahora ya podemos calcular $(A - I)^2$ o A^2 y efectuar las operaciones indicadas en (7.39) o (7.40) para escribir el resultado como una matriz 3×3.

$$e^{tA} = \begin{bmatrix} -2te^{t} + e^{2t} & (3t + 2)e^{t} - 2e^{2t} & -(t + 1)e^{t} + e^{2t} \\ -2(t + 1)e^{t} + 2e^{2t} & (3t + 5)e^{t} - 4e^{2t} & -(t + 2)e^{t} + 2e^{2t} \\ -2(t + 2)e^{t} + 4e^{2t} & (3t + 8)e^{t} - 8e^{2t} & -(t + 4)e^{t} + 4e^{2t} \end{bmatrix}.$$

7.15 Ejercicios

Para cada una de las matrices de los ejercicios 1 al 6, expresar e^{tA} como un polinomio en A.

1. $A = \begin{bmatrix} 5 & -2 \\ 4 & -1 \end{bmatrix}$. 2. $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$. 3. $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$.

4. $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$. 5. $A = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$. 6. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$.

7. a) Se sabe que una matriz A, 3×3, tiene todos sus autovalores iguales a λ. Demostrar que

$$e^{tA} = \tfrac{1}{2}e^{\lambda t}\{(\lambda^2t^2 - 2\lambda t + 2)I + (-2\lambda t^2 + 2t)A + t^2A^2\}.$$

b) Encontrar una fórmula análoga si A es una matriz 4×4 con todos sus autovalores iguales a λ.

En cada uno de los ejercicios del 8 al 15, resolver el sistema $Y' = AY$ sujeto a la condición inicial dada.

8. $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$, $Y(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$. 9. $A = \begin{bmatrix} -5 & 3 \\ -15 & 7 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

10. $A = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$. 11. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, $Y(0) = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$.

12. $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. 13. $A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix}$.

14. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. 15. $A = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}$.

16. Este ejercicio esboza una demostración de la ecuación (7.38) utilizada en la demostración del teorema 7.11. Sea $L_k(\lambda)$ el polinomio en λ de grado $n-1$ definido mediante

$$L_k(\lambda) = \prod_{\substack{j=1 \\ j\neq k}}^{n} \frac{\lambda - \lambda_j}{\lambda_k - \lambda_j},$$

donde $\lambda_1, \ldots, \lambda_n$ son n escalares distintos,

a) Demostrar que

$$L_k(\lambda_i) = \begin{cases} 0 & \text{si} \quad \lambda_i \neq \lambda_k, \\ 1 & \text{si} \quad \lambda_i = \lambda_k. \end{cases}$$

b) Sean $y_1, \ldots, y_n$, n escalares arbitrarios, y sea

$$p(\lambda) = \sum_{k=1}^{n} y_k L_k(\lambda).$$

Demostrar que $p(\lambda)$ es el único polinomio de grado $\leq n-1$ que satisface las n ecuaciones

$$p(\lambda_k) = y_k \quad \text{para} \quad k = 1, 2, \ldots, n.$$

c) Demostrar que $\sum_{k=1}^{n} L_k(\lambda) = 1$ para todo λ, y deducir que para toda matriz cuadrada A tenemos

$$\sum_{k=1}^{n} L_k(A) = I,$$

donde I es la matriz identidad.

7.16 Sistemas lineales no homogéneos con coeficientes constantes

A continuación consideramos el problema de valor inicial no homogéneo

$$(7.41) \qquad Y'(t) = AY(t) + Q(t), \qquad Y(a) = B,$$

en un intervalo J. A es una matriz constante $n \times n$, Q es una función vectorial n-dimensional (considerado como una matriz $n \times 1$) continua en J, y a es un punto dado en J. Podemos obtener una fórmula explícita para la solución de este problema siguiendo el mismo proceso que en el caso escalar.

Multiplicamos primero ambos miembros de (7.41) por la exponencial matricial e^{-tA} y pongamos la ecuación diferencial en la forma

(7.42) $$e^{-tA}\{Y'(t) - AY(t)\} = e^{-tA}Q(t).$$

El primer miembro de (7.42) es la derivada del producto $e^{-tA}Y(t)$. Por consiguiente, si integramos ambos miembros de (7.42) entre a y x, donde $x \in J$, obtenemos

$$e^{-xA}Y(x) - e^{-aA}Y(a) = \int_a^x e^{-tA}Q(t)\,dt.$$

Multiplicando por e^{xA} obtenemos la fórmula explícita (7.43) que aparece en el teorema siguiente.

TEOREMA 7.13. *Sean A una matriz constante $n\times n$ y Q una función vectorial n-dimensional continua en un intervalo J. Entonces el problema de valor inicial*

$$Y'(t) = AY(t) + Q(t), \qquad Y(a) = B,$$

tiene una solución única en J dada por la fórmula explícita

(7.43) $$Y(x) = e^{(x-a)A}B + e^{xA}\int_a^x e^{-tA}Q(t)\,dt.$$

Como en el caso homogéneo, la dificultad al aplicar esta fórmula en los casos prácticos radica en el cálculo de las exponenciales de matrices.

Obsérvese que el primer término, $e^{(x-a)A}B$, es la solución del problema homogéneo $Y'(t) = AY(t)$, $Y(a) = B$. El segundo término es la solución del problema no homogéneo

$$Y'(t) = AY(t) + Q(t), \qquad Y(a) = O.$$

Ilustramos el teorema 7.13 con un ejemplo.

EJEMPLO. Resolver el problema de valor inicial

$$Y'(t) = AY(t) + Q(t), \qquad Y(0) = B,$$

en el intervalo $(-\infty, +\infty)$, donde

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}, \qquad Q(t) = \begin{bmatrix} e^{2t} \\ 0 \\ te^{2t} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Solución. De acuerdo con el teorema 7.13, la solución viene dada por

$$(7.44) \qquad Y(x) = e^{xA}\int_0^x e^{-tA}Q(t)\,dt = \int_0^x e^{(x-t)A}Q(t)\,dt.$$

Los autovalores de A son 2, 2 y 4. Para calcular e^{xA} utilizamos la fórmula del caso III, sección 7.14, obteniendo

$$\begin{aligned} e^{xA} &= e^{2x}\{I + x(A-2I)\} + \tfrac{1}{4}(e^{4x} - e^{2x})(A-2I)^2 - \tfrac{1}{2}xe^{2x}(A-2I)^2 \\ &= e^{2x}\{I + x(A-2I) + \tfrac{1}{4}(e^{2x} - 2x - 1)(A-2I)^2\}. \end{aligned}$$

Podemos reemplazar x por $x - t$ en esta fórmula para obtener $e^{(x-t)A}$. Por consiguiente el integrando en (7.44) es

$$\begin{aligned} e^{(x-t)A}Q(t) &= e^{2(x-t)}\{I + (x-t)(A-2I) + \tfrac{1}{4}[e^{2(x-t)} - 2(x-t) - 1](A-2I)^2\}Q(t) \\ &= e^{2x}\begin{bmatrix} 1 \\ 0 \\ t \end{bmatrix} + (A-2I)e^{2x}\begin{bmatrix} x-t \\ 0 \\ t(x-t) \end{bmatrix} + \frac{1}{4}(A-2I)^2e^{2x}\begin{bmatrix} e^{2x}e^{-2t} - 2(x-t) - 1 \\ 0 \\ e^{2x}te^{-2t} - 2t(x-t) - t \end{bmatrix}. \end{aligned}$$

Integrando, encontramos

$$\begin{aligned} Y(x) = \int_0^x e^{(x-t)A}Q(t)\,dt &= e^{2x}\begin{bmatrix} x \\ 0 \\ \tfrac{1}{2}x^2 \end{bmatrix} + (A-2I)e^{2x}\begin{bmatrix} \tfrac{1}{2}x^2 \\ 0 \\ \tfrac{1}{6}x^3 \end{bmatrix} \\ &\quad + \frac{1}{4}(A-2I)^2e^{2x}\begin{bmatrix} \tfrac{1}{2}e^{2x} - \tfrac{1}{2} - x - x^2 \\ 0 \\ \tfrac{1}{4}e^{2x} - \tfrac{1}{4} - \tfrac{1}{2}x - \tfrac{1}{2}x^2 - \tfrac{1}{3}x^3 \end{bmatrix}. \end{aligned}$$

Puesto que tenemos

$$A - 2I = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \end{bmatrix} \qquad \text{y} \qquad (A-2I)^2 = \begin{bmatrix} 2 & 0 & 2 \\ -2 & 0 & -2 \\ 2 & 0 & 2 \end{bmatrix},$$

encontramos

$$Y(x) = e^{2x}\begin{bmatrix} x \\ 0 \\ \frac{1}{2}x^2 \end{bmatrix} + e^{2x}\begin{bmatrix} \frac{1}{6}x^3 \\ -\frac{1}{6}x^3 \\ x^2 + \frac{1}{6}x^3 \end{bmatrix} + e^{2x}\begin{bmatrix} \frac{3}{8}e^{2x} - \frac{3}{8} - \frac{3}{4}x - \frac{3}{4}x^2 - \frac{1}{6}x^3 \\ -\frac{3}{8}e^{2x} + \frac{3}{8} + \frac{3}{4}x + \frac{3}{4}x^2 + \frac{1}{6}x^3 \\ \frac{3}{8}e^{2x} - \frac{3}{8} - \frac{3}{4}x - \frac{3}{4}x^2 - \frac{1}{6}x^3 \end{bmatrix}$$

$$= e^{2x}\begin{bmatrix} \frac{3}{8}e^{2x} - \frac{3}{8} + \frac{1}{4}x - \frac{3}{4}x^2 \\ -\frac{3}{8}e^{2x} + \frac{3}{8} + \frac{3}{4}x + \frac{3}{4}x^2 \\ \frac{3}{8}e^{2x} - \frac{3}{8} - \frac{3}{4}x + \frac{3}{4}x^2 \end{bmatrix}.$$

Las filas de esta matriz son las funciones pedidas y_1, y_2, y_3.

7.17 Ejercicios

1. Sea Z una solución del sistema no homogéneo

$$Z'(t) = AZ(t) + Q(t),$$

en un intervalo J con valor inicial $Z(a)$. Demostrar que existe una sola solución del sistema homogéneo

$$Y'(t) = AY(t) + Q(t)$$

en J con valor inicial $Y(a)$ y que viene dado por la fórmula

$$Y(t) = Z(t) + e^{(t-a)A}\{Y(a) - Z(a)\}.$$

A menudo se dispone de otros métodos para determinar una solución particular $Z(t)$ que asemeja a la función dada $Q(t)$. Los ejercicios 2, 3, 5, y 7 indican tales métodos para $Q(t) = C$, $Q(t) = e^{\alpha t}C$, $Q(t) = t^m C$, y $Q(t) = (\cos \alpha t)C + (\operatorname{sen} \alpha t)D$, donde C y D son vectores constantes. Si la solución particular así obtenida $Z(t)$ no tiene el valor inicial pedido, modificamos $Z(t)$ como se indica en el ejercicio 1 para obtener otra solución $Y(t)$ con el valor inicial exigido.

2. a) Sea A una matriz constante $n \times n$, B y C sean vectores constantes n-dimensionales. Demostrar que la solución del sistema

$$Y'(t) = AY(t) + C, \qquad Y(a) = B,$$

en $(-\infty, +\infty)$ viene dada por la fórmula

$$Y(x) = e^{(x-a)A}B + \left(\int_0^{x-a} e^{uA}\, du\right)C.$$

b) Si A es no singular, demostrar que la integral del apartado a) tiene el valor $\{e^{(x-a)A} - I\}A^{-1}$.
c) Calcular $Y(x)$ en forma explícita cuando

$$A = \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad B = \begin{bmatrix} b \\ c \end{bmatrix}, \quad a = 0.$$

3. Sean A una matriz constante $n \times n$, y B y C vectores constantes n-dimensionales, y sea α un escalar dado.
a) Demostrar que el sistema no homogéneo $Z'(t) = AZ(t) + e^{\alpha t}C$ tiene una solución de la forma $Z(t) = e^{\alpha t}B$ si, y sólo si, $(\alpha I - A)B = C$.
b) Si α no es un autovalor de A, demostrar que el vector B siempre se puede elegir de modo que el sistema del apartado a) tenga una solución de la forma $Z(t) = e^{\alpha t}B$.
c) Si α no es un autovalor de A, demostrar que toda solución del sistema $Y'(t) = AY(t) + e^{\alpha t}C$ tiene la forma $Y(t) = e^{tA}(Y(0) - B) + e^{\alpha t}B$, donde $B = (\alpha I - A)^{-1}C$.
4. Utilizar el método sugerido por el ejercicio 3 para hallar una solución del sistema no homogéneo $Y'(t) = AY(t) + e^{2t}C$ con

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

5. Sean A una matriz constante $n \times n$, B y C vectores constantes n-dimensionales, y m un entero positivo.
a) Demostrar que el sistema no homogéneo $Y'(t) = AY(t) + t^m C$, $Y(0) = B$, tiene una solución particular de la forma

$$Y(t) = B_0 + tB_1 + t^2 B_2 + \cdots + t^m B_m,$$

en donde $B_0, B_1, \ldots, B_m$ son vectores constantes, si y sólo si

$$C = -\frac{1}{m!} A^{m+1} B.$$

Determinar los coeficientes $B_0, B_1, \ldots, B_m$ para una tal solución.
b) Si A es no singular, demostrar que el vector inicial B siempre puede elegirse de manera que el sistema del apartado a) tenga una solución de la forma indicada.
6. Considérese el sistema no homogéneo

$$y_1' = 3y_1 + y_2 + t^3$$

$$y_2' = 2y_1 + 2y_2 + t^3.$$

a) Encontrar una solución particular de la forma $Y(t) = B_0 + tB_1 + t^2B_2 + t^3B_3$.
b) Encontrar una solución del sistema con la condición $y_1(0) = y_2(0) = 1$.
7. Sean A una matriz constante $n \times n$, B, C, D vectores constantes n-dimensionales y α un número real no nulo dado. Demostrar que el sistema no homogéneo

$$Y'(t) = AY(t) + (\cos \alpha t)C + (\operatorname{sen} \alpha t)D, \qquad Y(0) = B,$$

tiene una solución particular de la forma

$$Y(t) = (\cos \alpha t)E + (\operatorname{sen} \alpha t)F,$$

siendo E y F vectores constantes, si y sólo si

$$(A^2 + \alpha^2 I)B = -(AC + \alpha D).$$

Determinar E y F en función de A, B, C para una tal solución. Obsérvese que si $A^2 + \alpha^2 I$ es no singular, el vector inicial B siempre se puede elegir de modo que el sistema tenga una solución de la forma indicada.

8. a) Hallar una solución particular del sistema no homogéneo

$$y_1' = y_1 + 3y_2 + 4 \operatorname{sen} 2t$$

$$y_2' = y_1 - y_2 .$$

b) Hallar una solución del sistema con las condiciones $y_1(0) = y_2(0) = 1$.

En cada uno de los ejercicios 9 al 12, resolver el sistema no homogéneo $Y'(t) = AY(t) + Q(t)$ sujeto a la condición inicial indicada.

9. $A = \begin{bmatrix} 4 & 1 \\ -2 & 1 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} 0 \\ -2e^t \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$

10. $A = \begin{bmatrix} -5 & -1 \\ 1 & -3 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} e \\ e^{2t} \end{bmatrix}, \quad Y(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$

11. $A = \begin{bmatrix} -5 & -1 \\ 2 & -3 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} 7e^t - 27 \\ -3e^t + 12 \end{bmatrix}, \quad Y(0) = \begin{bmatrix} -\frac{1007}{442} \\ \frac{707}{221} \end{bmatrix}.$

12. $A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} t^2 \\ 2t \\ t \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix}.$

7.18 Sistema lineal general $Y'(t) = P(t)Y(t) + Q(t)$

El teorema 7.13 da una fórmula explícita para la solución del sistema lineal

$$Y'(t) = AY(t) + Q(t), \qquad Y(a) = B,$$

donde A es una matriz constante $n \times n$ y $Q(t)$, $Y(t)$ son matrices columna $n \times 1$. Volvemos ahora al caso más general

$$(7.45) \qquad Y'(t) = P(t)Y(t) + Q(t), \qquad Y(a) = B,$$

en donde la matriz $P(t)$, $n \times n$, no es necesariamente constante.

Si P y Q son continuas en un intervalo abierto J, un teorema general de existencia y unicidad que demostraremos más adelante nos dice que para cada a de J y cada vector inicial B existe una sola solución para el problema de valor inicial (7.45). En esta sección utilizamos este resultado para obtener una fórmula para la solución, que generaliza el teorema 7.13.

En el caso escalar ($n = 1$) la ecuación diferencial (7.45) puede resolverse del modo siguiente. Sea $A(x) = \int_a^x P(t)dt$, multipliquemos entonces ambos miembros de (7.45) por $e^{-A(t)}$ poniendo la ecuación diferencial en la forma

$$(7.46) \qquad e^{-A(t)}\{Y'(t) - P(t)Y(t)\} = e^{-A(t)}Q(t).$$

El primer miembro es la derivada del producto $e^{-A(t)}Y(t)$. Por consiguiente, podemos integrar ambos miembros entre a y x, siendo a y x puntos de J, y obtenemos

$$e^{-A(x)}Y(x) - e^{-A(a)}Y(a) = \int_a^x e^{-A(t)}Q(t)\,dt.$$

Multiplicando por $e^{A(x)}$ se llega a la fórmula explícita

$$(7.47) \qquad Y(x) = e^{A(x)}e^{-A(a)}Y(a) + e^{A(x)} \int_a^x e^{-A(t)}Q(t)\,dt.$$

La única parte de este razonamiento que no es aplicable a las funciones matriciales es la afirmación de que el primer miembro de (7.46) es la derivada del producto $e^{-A(t)}Y(t)$. En este punto utilizamos el hecho de que la derivada de $e^{-A(t)}$ es $-P(t)e^{-A(t)}$. En el caso escalar esto es una consecuencia de la fórmula siguiente para la derivación de las funciones exponenciales

$$\text{Si} \qquad E(t) = e^{A(t)}, \quad \text{entonces} \quad E'(t) = A'(t)e^{A(t)}.$$

Desgraciadamente, esta fórmula de derivación no siempre es válida cuando A es una función matricial. Por ejemplo, es falsa para la función matricial 2×2 $A(t) = \begin{bmatrix} 1 & t \\ 0 & 0 \end{bmatrix}$ (Ver ejercicio 7 de la sección 7.12). Por tanto es necesario una

modificación del razonamiento para extender la ecuación (7.47) al caso de las matrices.

Supongamos que multiplicamos cada miembro de (7.45) por una matriz $F(t)$, $n \times n$, no especificada. Esto nos da la relación

$$F(t)Y'(t) = F(t)P(t)Y(t) + F(t)Q(t).$$

Sumemos ahora a los dos miembros $F'(t)Y(t)$ con la idea de transformar el primer miembro en la derivada del producto $F(t)Y(t)$. Si hacemos esto, la última ecuación nos da

$$\{F(t)Y(t)\}' = \{F'(t) + F(t)P(t)\}Y(t) + F(t)Q(t).$$

Si podemos elegir la matriz $F(t)$ de manera que la suma $\{F'(t) + F(t)P(t)\}$ del segundo miembro sea la matriz cero, la última ecuación se simplifica y queda

$$\{F(t)Y(t)\}' = F(t)Q(t).$$

Integrando ésta desde a a x obtenemos

$$F(x)Y(x) - F(a)Y(a) = \int_a^x F(t)Q(t)\,dt.$$

Si, además, la matriz $F(x)$ es no singular, obtenemos la fórmula explícita

(7.48) $$Y(x) = F(x)^{-1}F(a)Y(a) + F(x)^{-1}\int_a^x F(t)Q(t)\,dt.$$

Ésta es una generalización de la fórmula escalar (7.47). El proceso funcionará bien si podemos encontrar una función matricial $F(t)$, $n \times n$, que satisfaga la ecuación diferencial matricial

$$F'(t) = -F(t)P(t)$$

y que no sea singular.

Obsérvese que esta ecuación diferencial es muy parecida a la ecuación diferencial original (7.45) con $Q(t) = 0$, salvo que la función incógnita $F(t)$ es una matriz cuadrada en lugar de una matriz columna. Asimismo, la función incógnita está multiplicada a la derecha por $-P(t)$ en lugar de estarlo a la izquierda por $P(t)$.

Seguidamente demostraremos que la ecuación diferencial para F siempre tiene una solución no singular. La demostración dependerá del siguiente teorema de existencia para sistemas lineales homogéneos.

TEOREMA 7.14. *Supongamos que $A(t)$ es una función matricial $n \times n$ continua en un intervalo abierto J. Si $a \in J$ y B es un vector n-dimensional dado, el sistema lineal homogéneo*

$$Y'(t) = A(t)Y(t), \qquad Y(a) = B,$$

tiene en J una solución Y que es un vector n-dimensional

En la sección 7.21 damos una demostración de este teorema. Con la ayuda de éste podemos demostrar el siguiente.

TEOREMA 7.15. *Dada una función matricial P, $n \times n$, continua en un intervalo abierto J, y dado un punto cualquiera a en J, existe una función matricial F, $n \times n$, que satisface la ecuación diferencial matricial*

$$F'(x) = -F(x)P(x) \tag{7.49}$$

en J con el valor inicial $F(a) = I$. Además, $F(x)$ es no singular para cada x en J.

Demostración. Sea $Y_k(x)$ una solución vectorial de la ecuación diferencial

$$Y_k'(x) = -P(x)^t Y_k(x)$$

en J con vector inicial $Y_k(a) = I_k$, donde I_k es la columna k de la matriz identidad I, $n \times n$. Aquí $P(x)^t$ representa la transpuesta de $P(x)$. Sea $G(x)$ la matriz $n \times n$ cuya columna k es $Y_k(x)$. Entonces G satisface la ecuación diferencial matricial

$$G'(x) = -P(x)^t G(x) \tag{7.50}$$

en J con valor inicial $G(a) = I$. Tomemos ahora la transpuesta de cada miembro de (7.50). Puesto que la transpuesta de un producto es el producto de las transpuestas en orden invertido, obtenemos

$$\{G'(x)\}^t = -G(x)^t P(x).$$

También, la transpuesta de la derivada G' es la derivada de la transpuesta G^t. Por tanto la matriz $F(x) = G(x)^t$ satisface la ecuación diferencial (7.49) con valor inicial $F(a) = I$.

Demostremos ahora que $F(x)$ es no singular construyendo su inversa. Sea H la función matricial $n \times n$ cuya columna k es la solución de la ecuación diferencial

$$Y'(x) = P(x)Y(x)$$

con valor inicial $Y(a) = I_k$, la columna k de I. Entonces H satisface el problema de valor inicial

$$H'(x) = P(x)H(x), \qquad H(a) = I,$$

en J. El producto $F(x)H(x)$ tiene derivada

$$F(x)H'(x) + F'(x)H(x) = F(x)P(x)H(x) - F(x)P(x)H(x) = O$$

para cada x de J. Por consiguiente el producto $F(x)H(x)$ es constante, $F(x)H(x) = F(a)H(a) = I$, así que $H(x)$ es la inversa de $F(x)$. Esto completa la demostración.

En el teorema que sigue quedan resumidos los resultados de esta sección.

TEOREMA 7.16. *Dadas una función matricial P, $n\times n$, y una función vectorial n-dimensional Q, ambas continuas en un intervalo abierto J, la solución del problema de valor inicial*

$$Y'(x) = P(x)Y(x) + Q(x), \qquad Y(a) = B, \tag{7.51}$$

en J viene dado por la fórmula

$$Y(x) = F(x)^{-1}Y(a) + F(x)^{-1}\int_a^x F(t)Q(t)\,dt. \tag{7.52}$$

La matriz $F(x)$, $n\times n$, es la transpuesta de la matriz cuya columna k es la solución del problema de valor inicial

$$Y'(x) = -P(x)^t Y(x), \qquad Y(a) = I_k, \tag{7.53}$$

en donde I_k es la columna k de la matriz identidad I.

Aunque el teorema 7.16 proporciona una fórmula explícita para la resolución del sistema lineal general (7.51), tal fórmula no siempre resulta útil para calcular la solución debido a la dificultad que supone la determinación de la función matricial F. Ello exige la solución de n sistemas lineales homogéneos (7.53). En la siguiente sección se expone un método con series de potencias que a veces se emplea en la resolución de sistemas lineales homogéneos.

Una vez más recordamos al lector que la demostración del teorema 7.16 se basó en el teorema de existencia 7.14, para los sistemas lineales homogéneos, que todavía no hemos demostrado.

7.19 Resolución de sistemas lineales homogéneos mediante series de potencias

Consideremos un sistema lineal homogéneo

$$Y'(x) = A(x)Y(x), \qquad Y(0) = B, \tag{7.54}$$

en el que la matriz A, $n \times n$, dada tiene un desarrollo en serie de potencias de x convergente en un cierto intervalo abierto que contenga el origen, tal como

$$A(x) = A_0 + xA_1 + x^2A_2 + \cdots + x^kA_k + \ldots, \qquad \text{para } |x| < r_1,$$

donde los coeficientes $A_0, A_1, A_2, \ldots$ son matrices $n \times n$ dadas. Intentemos encontrar una serie de potencias solución de la forma

$$Y(x) = B_0 + xB_1 + x^2B_2 + \cdots + x^kB_k + \cdots,$$

con los vectores $B_0, B_1, B_2, \ldots$ como coeficientes. Puesto que $Y(0) = B_0$, la condición inicial se satisfará tomando $B_0 = B$, el vector inicial prescrito. Para determinar los coeficientes restantes sustituimos la serie de potencias que da $Y(x)$ en la ecuación diferencial e igualamos los coeficientes de potencias semejantes de x obteniendo el siguiente sistema de ecuaciones

$$B_1 = A_0B, \qquad (k+1)B_{k+1} = \sum_{r=0}^{k} A_rB_{k-r} \qquad \text{para } k = 1, 2, \ldots. \tag{7.55}$$

Estas ecuaciones pueden resolverse en forma sucesiva respecto a los vectores $B_1, B_2, \ldots.$ Si la serie de potencias de $Y(x)$ que resulta converge en un cierto intervalo $|x| < r_2$, entonces $Y(x)$ será una solución del problema de valor inicial (7.54) en el intervalo $|x| < r$, siendo $r = \min\{r_1, r_2\}$.

Por ejemplo, si $A(x)$ es una matriz constante A, entonces $A_0 = A$ y $A_k = O$ para $k \geq 1$, así que el sistema de ecuaciones (7.55) se convierte en

$$B_1 = AB, \qquad (k+1)B_{k+1} = AB_k \quad \text{para } k \geq 1.$$

Resolviendo esas ecuaciones sucesivamente encontramos

$$B_k = \frac{1}{k!}A^kB \quad \text{para } k \geq 1.$$

Por tanto en este caso la serie solución toma la forma

$$Y(x) = B + \sum_{k=1}^{\infty} \frac{x^k}{k!}A^kB = e^{xA}B.$$

Esto está de acuerdo con el resultado obtenido antes para los sistemas lineales homogéneos con coeficientes constantes.

7.20 Ejercicios

1. Sean p una función de valores reales y Q una función matricial $n\times 1$, ambas continuas en un intervalo J. Sea A una matriz constante $n\times n$. Demostrar que el problema de valor inicial

$$Y'(x) = p(x)AY(x) + Q(x), \qquad Y(a) = B,$$

tiene la solución

$$Y(x) = e^{q(x)A}B + e^{q(x)A}\int_a^x e^{-q(t)A}Q(t)\,dt$$

en J, donde $q(x) = \int_a^x p(t)\,dt$.

2. Considerar el caso particular del ejercicio 1 en el que A sea no singular, $a=0$, $p(x)=2x$, y $Q(x) = xC$, siendo C un vector constante. Demostrar que la solución se convierte en

$$Y(x) = e^{x^2A}(B + \tfrac{1}{2}A^{-1}C) - \tfrac{1}{2}A^{-1}C.$$

3. Sean $A(t)$ una función matricial $n\times n$ y $E(t) = e^{A(t)}$. Sean $Q(t)$, $Y(t)$ y B matrices columna $n\times 1$. Supongamos que

$$E'(t) = A'(t)E(t)$$

en un intervalo abierto J. Si $a \in J$ y si A' y Q son continuas en J, demostrar que el problema de valor inicial

$$Y'(t) = A'(t)Y(t) + Q(t), \qquad Y(a) = B,$$

tiene la siguiente solución en J:

$$Y(x) = e^{A(x)}e^{-A(a)}B + e^{A(x)}\int_a^x e^{-A(t)}Q(t)\,dt.$$

4. Sea $E(t) = e^{A(t)}$. Este ejercicio muestra ejemplos de funciones matriciales $A(t)$ para las que $E'(t) = A'(t)E(t)$.
a) Que sea $A(t) = t^rA$, donde A es una matriz constante $n\times n$ y r un entero positivo. Demostrar que $E'(t) = A'(t)E(t)$ en $(-\infty, +\infty)$.
b) Sea $A(t)$ un polinomio en t con matrices como coeficientes

$$A(t) = \sum_{r=0}^{m} t^r A_r,$$

siendo los coeficientes permutables $A_rA_s = A_sA_r$ para todos los r y s. Demostrar que $E'(t) = A'(t)E(t)$ en $(-\infty, +\infty)$.
c) Resolver el sistema lineal homogéneo

$$Y'(t) = (I + tA)Y(t), \qquad Y(0) = B$$

en el intervalo $(-\infty, \infty)$, donde A es una matriz constante $n \times n$.

5. Supongamos que la función matricial, $n \times n$, $A(x)$ admite un desarrollo en serie de potencias convergente para $|x| < r$. Desarrollar un procedimiento para resolver, mediante desarrollo en serie, el siguiente sistema lineal homogéneo de segundo orden:

$$Y''(x) = A(x)Y(x), \qquad \text{con} \quad Y(0) = B, \qquad Y'(0) = C.$$

6. Considerar el sistema de segundo orden $Y''(x) + AY(x) = 0$, con $Y(0) = B$, $Y'(0) = C$ donde A es una matriz constante $n \times n$. Demostrar que el sistema tiene la serie de potencias solución

$$Y(x) = \left(I + \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k} A^k}{(2k)!}\right) B + \left(xI + \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k+1} A^k}{(2k+1)!}\right) C$$

convergente para $-\infty < x < +\infty$.

7.21 Demostración del teorema de existencia por el método de las aproximaciones sucesivas

En esta sección demostramos la existencia y unicidad de una solución para cualquier sistema lineal homogéneo

$$Y'(t) = A(t)Y(t), \tag{7.56}$$

donde $A(t)$ es una función matricial $n \times n$, continua en un intervalo abierto J. Demostraremos que para un punto cualquiera a de J y cualquier vector inicial B existe exactamente una solución $Y(t)$ en J que satisface la condición inicial $Y(a) = B$.

Usaremos el *método de aproximaciones sucesivas*, que es un método iterativo aplicable también a otros muchos problemas. Fue publicado primero por Liouville en 1838 en conexión con el estudio de las ecuaciones diferenciales lineales de segundo orden. Más tarde fue extendido por J. Caqué en 1864, L. Fuchs en 1870, y G. Peano en 1888 al estudio de las ecuaciones lineales de orden n. En 1890 Émile Picard (1856-1941) extendió el método al abarcar también las ecuaciones diferenciales no lineales. En reconocimiento a su contribución fundamental, algunos autores lo citan como el *método de Picard*. El interés del método no es

sólo teórico sino que también se usa en algunos casos para obtener las aproximaciones numéricas de las soluciones.

El método empieza tomando una supuesta solución de la ecuación (7.56). Tomamos como tal el vector inicial dado B, si bien esto no es esencial. Sustituimos entonces esta supuesta solución en el segundo miembro de la ecuación y obtenemos una nueva ecuación diferencial

$$Y'(t) = A(t)B .$$

En esta ecuación el segundo miembro ya no contiene la función incógnita, así que la ecuación puede resolverse inmediatamente integrando ambos miembros entre a y x, siendo x un punto arbitrario de J. Esta ecuación tiene exactamente una solución Y_1 en J que satisface la condición inicial $Y_1(a) = B$, o sea

$$Y_1(x) = B + \int_a^x A(t)B\,dt .$$

Reemplazamos ahora $Y(t)$ por $Y_1(t)$ en el segundo miembro de la ecuación diferencial original (7.56) obteniendo una nueva ecuación diferencial

$$Y'(t) = A(t)Y_1(t) .$$

Esta ecuación tiene una sola solución Y_2 en J con la condición $Y_2(a) = B$,

$$Y_2(x) = B + \int_a^x A(t)Y_1(t)\,dt . \tag{7.57}$$

Sustituimos entonces Y_2 en el segundo miembro de (7.56) y resolvemos la ecuación resultante para determinar Y_3 con $Y_3(a) = B$, y así sucesivamente. Este proceso engendra una sucesión de funciones $Y_0, Y_1, Y_2, \ldots$, en la que $Y_0 = B$ y donde Y_{k+1} se determina a partir de Y_k mediante la fórmula de recurrencia

$$Y_{k+1}(x) = B + \int_a^x A(t)Y_k(t)\,dt \qquad \text{para} \quad k = 0, 1, 2, \ldots . \tag{7.58}$$

Nuestro objetivo es demostrar que la sucesión de funciones así definida converge hacia una función límite Y que es una solución de la ecuación diferencial (7.56) en J y que también satisface la condición inicial $Y(a) = B$. Las funciones $Y_0, Y_1, Y_2, \ldots$ se llaman *aproximaciones sucesivas* de Y. Antes de estudiar la convergencia del proceso ilustramos el método con un ejemplo.

EJEMPLO. Consideremos el problema de valor inicial $Y'(t) = AY(t)$, $Y(0) = B$,

donde A es una matriz constante $n \times n$. Sabemos que la solución está dada por la fórmula $Y(x) = e^{xA}B$ para todo x real. Demostraremos cómo esta solución se puede obtener por el método de las aproximaciones sucesivas.

La hipótesis inicial es $Y_0(x) = B$. La fórmula de recurrencia (7.58) nos da

$$Y_1(x) = B + \int_0^x AB\, dt = B + xAB,$$

$$Y_2(x) = B + \int_0^x AY_1(t)\, dt = B + \int_0^x (AB + tA^2B)\, dt = B + xAB + \tfrac{1}{2}x^2A^2B.$$

Por inducción encontramos

$$Y_k(x) = B + xAB + \tfrac{1}{2}x^2A^2B + \cdots + \frac{1}{k!}x^kA^kB = \left(\sum_{r=0}^{k} \frac{(xA)^r}{r!}\right)B.$$

La suma del segundo miembro es una suma parcial de la serie de e^{xA}. Por consiguiente cuando $k \to \infty$ encontramos

$$\lim_{k\to\infty} Y_k(x) = e^{xA}B$$

para todo x. Así pues, en este ejemplo podemos demostrar directamente que las aproximaciones sucesivas convergen hacia una solución del problema de valor inicial en $(-\infty, +\infty)$.

Demostración de la convergencia de la sucesión de las aproximaciones sucesivas. Volvamos ahora a la sucesión general definida por la fórmula de recurrencia (7.58). Para demostrar que la sucesión converge escribimos cada término $Y_k(x)$ cómo una suma telescópica,

$$(7.59) \qquad Y_k(x) = Y_0(x) + \sum_{m=0}^{k-1} \{Y_{m+1}(x) - Y_m(x)\}.$$

Para demostrar que $Y_k(x)$ tiende hacia un límite cuando $k \to \infty$ demostraremos que la serie

$$(7.60) \qquad \sum_{m=0}^{\infty} \{Y_{m+1}(x) - Y_m(x)\}$$

converge para cada x de J. A tal fin basta demostrar que la serie

$$(7.61) \qquad \sum_{m=0}^{\infty} \|Y_{m+1}(x) - Y_m(x)\|$$

converge. En esta serie utilizamos el concepto de norma de una matriz introducido en la sección 7.3; la norma de una matriz es la suma de los valores absolutos de todos sus términos.

Consideremos un subintervalo acotado y cerrado J_1 de J que contenga a. Demostraremos que para todo x de J_1 la serie (7.61) es mayorada por una serie convergente de constantes independientes de x. Esto implica que la serie converge *uniformemente* en J_1.

Para hacer una estimación de la magnitud de los términos de (7.61) utilizamos repetidamente la fórmula de recurrencia. Inicialmente, tenemos

$$Y_1(x) - Y_0(x) = \int_a^x A(t)B\,dt\,.$$

Para mayor sencillez, supongamos que $a < x$. Entonces podemos escribir

$$\|Y_1(x) - Y_0(x)\| = \left\| \int_a^x A(t)B\,dt \right\| \leq \int_a^x \|A(t)\|\ \|B\|\,dt\,. \tag{7.62}$$

Puesto que cada elemento de $A(t)$ es función continua en J, cada elemento está acotado en el intervalo cerrado acotado J_1. Por consiguiente $\|A(t)\| \leq M$, donde M es la suma de las cotas de todos los elementos de $A(t)$ en el intervalo J_1. (El número M depende de J_1.) Por consiguiente el integrando de (7.62) está acotado por $\|B\|M$, así que tenemos

$$\|Y_1(x) - Y_0(x)\| \leq \int_a^x \|B\|\,M\,dt = \|B\|\,M(x-a)$$

para todo $x > a$ en J_1.

Utilicemos ahora la fórmula de recurrencia una vez más para expresar la diferencia $Y_2 - Y_1$ en función de $Y_1 - Y_0$, y luego empleamos la estimación que acabamos de obtener para $Y_1 - Y_0$, y se deduce

$$\|Y_2(x) - Y_1(x)\| = \left\| \int_a^x A(t)\{Y_1(t) - Y_0(t)\}\,dt \right\| \leq \int_a^x \|A(t)\|\ \|B\|\,M(t-a)\,dt$$
$$\leq \|B\|\,M^2 \int_a^x (t-a)\,dt = \|B\|\,\frac{M^2(x-a)^2}{2!}$$

para todo $x > a$ en J_1. Por inducción encontramos

$$\|Y_{m+1}(x) - Y_m(x)\| \leq \|B\|\,\frac{M^{m+1}(x-a)^{m+1}}{(m+1)!} \qquad \text{para} \quad m = 0, 1, 2, \ldots,$$

y para todo $x > a$ de J_1. Si $x < a$ un razonamiento parecido da la misma desigualdad con $|x - a|$ en lugar de $(x - a)$. Si designamos con L la longitud del intervalo

J_1, entonces tenemos $|x - a| \leq L$ para todo x de J_1 así que obtenemos la acotación

$$\|Y_{m+1}(x) - Y_m(x)\| \leq \|B\| \frac{M^{m+1}L^{m+1}}{(m+1)!} \quad \text{para} \quad m = 0, 1, 2, \ldots,$$

para todo x de J_1. Por consiguiente la serie de (7.61) es mayorada por la serie convergente

$$\|B\| \sum_{m=0}^{\infty} \frac{(ML)^{m+1}}{(m+1)!} = \|B\| (e^{ML} - 1).$$

Esto demuestra que la serie (7.61) converge uniformemente en J_1.

El razonamiento anterior demuestra que la sucesión de aproximaciones sucesivas siempre converge y en J_1 la convergencia es uniforme. Sea Y la función límite. Esto es, definimos $Y(x)$ para cada x en J_1 por

$$Y(x) = \lim_{k \to \infty} Y_k(x).$$

Demostraremos que Y tiene las propiedades siguientes:

a) Y es continua en J_1.

b) $Y(x) = B + \int_a^x A(t)\, Y(t)\, dt$ para todo x de J_1.

c) $Y(a) = B$ e $Y'(x) = A(x)\, Y(x)$ para todo x de J_1.

La parte c) demuestra que Y es una solución del problema de valor inicial en J_1.

Demostración de a). Cada función Y_k es una matriz columna cuyos elementos son funciones escalares, continuas en J_1. Cada elemento de la función límite Y es el límite de una sucesión uniformemente convergente de funciones continuas, así que, según el teorema 11.1 del volumen I, cada elemento de Y es también continua en J_1. Por consiguiente Y es continua en J_1.

Demostración de b). La fórmula de recurrencia (7.58) establece que

$$Y_{k+1}(x) = B + \int_a^x A(t) Y_k(t)\, dt.$$

Por consiguiente

$$Y(x) = \lim_{k \to \infty} Y_{k+1}(x) = B + \lim_{k \to \infty} \int_a^x A(t) Y_k(t)\, dt = B + \int_a^x A(t) \lim_{k \to \infty} Y_k(t)\, dt$$
$$= B + \int_a^x A(t) Y(t)\, dt.$$

El intercambio del símbolo de límite con el signo integral es válido por la convergencia uniforme de la sucesión $\{Y_k\}$ en J_1.

Demostración de c). La ecuación $Y(a) = B$ rèsulta al momento de b). Debido al apartado a), el integrando de b) es continuo en J_1 así que, en virtud del primer teorema fundamental del cálculo, $Y'(x)$ existe y es igual a $A(x)\,Y(x)$ en J_1.

El intervalo J_1 sabemos que es un subintervalo cualquiera de J cerrado y acotado y que contenga a. Si J_1 se amplía, el proceso para obtener $Y(x)$ no varía porque sólo incluye la integración de a a x. Puesto que para todo x de J existe un subintervalo de J cerrado y acotado y que contiene a y x, existe una solución en todo el intervalo J.

TEOREMA 7.17. TEOREMA DE UNICIDAD PARA SISTEMAS LINEALES HOMOGÉNEOS. *Si $A(t)$ es continua en un intervalo abierto J, la ecuación diferencial*

$$Y'(t) = A(t)Y(t)$$

tiene a lo más una solución en J que satisface una condición inicial dada $Y(a)=B$.

Demostración. Sean Z e Y dos soluciones en J. Sea J_1 cualquier subintervalo de J acotado y cerrado que contenga a. Demostraremos que $Z(x) = Y(x)$ para todo x en J_1. Esto implica que $Z = Y$ en todo el intervalo J.

Puesto que Z e Y son soluciones tenemos

$$Z'(t) - Y'(t) = A(t)\{Z(t) - Y(t)\}.$$

Elijamos x en J_1 e integremos esta ecuación entre a y x con lo que obtenemos

$$Z(x) - Y(x) = \int_a^x A(t)\{Z(t) - Y(t)\}\,dt.$$

Ésta implica la desigualdad

$$\|Z(x) - Y(x)\| \le M\left|\int_a^x \|Z(t) - Y(t)\|\,dt\right|, \tag{7.63}$$

donde M es una cota superior de $\|A(t)\|$ en J_1. Sea M_1 una cota superior para la función continua $\|Z(t) - Y(t)\|$ en J_1. Entonces la desigualdad (7.63) nos da

$$\|Z(x) - Y(x)\| \le MM_1\,|x - a|. \tag{7.64}$$

Aplicando (7.64) en el segundo miembro de (7.63) obtenemos

$$\|Z(x) - Y(x)\| \le M^2M_1\left|\int_a^x |t - a|\,dt\right| = M^2M_1\frac{|x - a|^2}{2}.$$

Por inducción resulta

$$\|Z(x) - Y(x)\| \leq M^m M_1 \frac{|x-a|^m}{m!}. \tag{7.65}$$

Cuando $m \to \infty$ el segundo miembro tiende a 0, así que $Z(x) = Y(x)$. Esto completa la demostración.

Los resultados de esta sección pueden resumirse en el siguiente teorema de existencia y unicidad.

TEOREMA 7.18. *Sea A una función matricial $n \times n$ continua en un intervalo abierto J. Si $a \in J$ y B es un vector n-dimensional cualquiera, el sistema lineal homogéneo*

$$Y'(t) = A(t)Y(t), \qquad Y(a) = B,$$

tiene una y sólo una solución que es un vector n-dimensional en J.

7.22 Aplicación del método de aproximaciones sucesivas a los sistemas no lineales de primer orden.

El método de aproximaciones sucesivas también puede aplicarse a algunos sistemas no lineales. Consideremos un sistema de primer orden de la forma

$$Y' = F(t, Y), \tag{7.66}$$

donde F es una función vectorial n-dimensional dada, e Y es una función vectorial n-dimensional incógnita que hay que determinar. Queremos una solución Y que satisfaga la ecuación

$$Y'(t) = F[t, Y(t)]$$

para cada t en un cierto intervalo J y que también satisfaga una condición inicial dada, por ejemplo $Y(a) = B$, donde $a \in J$ y B es un vector n-dimensional dado.

Al igual que en el caso lineal, construimos una sucesión de aproximaciones sucesivas $Y_0, Y_1, Y_2 \ldots$, tomando $Y_0 = B$ y definiendo Y_{k+1} en función de Y_k mediante la fórmula de recurrencia

$$Y_{k+1}(x) = B + \int_a^x F[t, Y_k(t)]\,dt \quad \text{para} \quad k = 0, 1, 2, \ldots. \tag{7.67}$$

Imponiendo ciertas condiciones a F, esta sucesión convergerá hacia una función límite Y que satisfará la ecuación diferencial y la condición inicial dadas.

Antes de averiguar la convergencia del proceso vamos a discutir algunos ejemplos unidimensionales elegidos para ilustrar algunas de las dificultades que pueden presentarse en la práctica.

EJEMPLO 1. Consideremos el problema de valor inicial no lineal $y' = x^2 + y^2$ con $y = 0$ cuando $x = 0$. Calcularemos algunas aproximaciones de la solución. Elijamos $Y_0(x) = 0$ y determinaremos las tres siguientes aproximaciones como sigue:

$$Y_1(x) = \int_0^x t^2\,dt = \frac{x^3}{3},$$

$$Y_2(x) = \int_0^x [t^2 + Y_1^2(t)]\,dt = \int_0^x \left(t^2 + \frac{t^6}{9}\right)dt = \frac{x^3}{3} + \frac{x^7}{63},$$

$$Y_3(x) = \int_0^x \left[t^2 + \left(\frac{t^3}{3} + \frac{t^7}{63}\right)^2\right]dt = \frac{x^3}{3} + \frac{x^7}{63} + \frac{2x^{11}}{2079} + \frac{x^{15}}{59535}.$$

Ahora se ve el enorme trabajo que será necesario para calcular más aproximaciones. Por ejemplo, las dos aproximaciones siguientes Y_4 e Y_5 serán polinomios de grados 31 y 63, respectivamente.

El próximo ejemplo muestra otra dificultad que puede presentarse en el cálculo de las aproximaciones sucesivas.

EJEMPLO 2. Consideremos el problema, de valor inicial, no lineal $y' = 2x + e^y$, con $y = 0$ cuando $x = 0$. Comenzamos con la función inicial $Y_0(x) = 0$ y encontramos

$$Y_1(x) = \int_0^x (2t + 1)\,dt = x^2 + x,$$

$$Y_2(x) = \int_0^x (2t + e^{t^2+t})\,dt = x^2 + \int_0^x e^{t^2+t}\,dt.$$

Aquí, no se puede seguir adelante por el hecho de que la última integral no puede calcularse como expresión de funciones elementales. No obstante, para una x dada es posible calcular una aproximación numérica de la integral y con ello se obtiene una aproximación de $Y_2(x)$.

A causa de las dificultades que hemos visto en los dos últimos ejemplos, el método de aproximaciones sucesivas en la práctica no es el más útil para la determinación explícita de soluciones. El auténtico valor del método consiste en su empleo para establecer teoremas de existencia.

7.23 Demostración de un teorema de existencia y unicidad para sistemas no lineales de primer orden

Volvamos ahora al teorema de existencia y unicidad para sistemas no lineales de primer orden. Imponiendo restricciones adecuadas a la función que aparece en el segundo miembro de la ecuación diferencial

$$Y' = F(x, Y),$$

podemos extender el método de demostración usado en el caso lineal de la sección 7.21.

Sea J el intervalo abierto en el que deseamos una solución. Supongamos que $a \in J$ y que B es un vector n-dimensional dado. Designemos con S un conjunto en el espacio de dimensión $(n + 1)$ dado por

$$S = \{(x,Y) \mid \; |x - a| \leq h, \; \|Y - B\| \leq k\},$$

donde $h > 0$ y $k > 0$. [Si $n = 1$, es un rectángulo con centro en (a, B), con base $2h$ y altura $2k$.] Supongamos que el dominio de F incluye un conjunto S de este tipo y que F es acotada en S, sea

$$\|F(x, Y)\| \leq M \tag{7.68}$$

para todo punto (x, Y) de S, siendo M una constante positiva.

Seguidamente, supongamos que la función compuesta $G(x) = F(x, Y(x))$ es continua en el intervalo $(a - h, a + h)$ para toda función Y que sea continua en $(a - h, a + h)$ y que tenga la propiedad de que $(x, Y(x)) \in S$ para todo x en $(a - h, a + h)$. Esta hipótesis asegura la existencia de las integrales que se presentan en el método de aproximaciones sucesivas, e implica también la continuidad de las funciones así construidas.

Finalmente, supongamos que F satisface una condición de la forma

$$\|F(x, Y) - F(x, Z)\| \leq A \, \|Y - Z\|$$

para todo par de puntos (x, Y) y (x, Z) de S, siendo A una constante positiva. Ésta se llama una *condición de Lipschitz* en honor de Rudolph Lipschitz, que fue el primero en introducirla en 1876. Una tal condición no es un restricción muy seria para una función y nos permite extender la demostración de la existencia y unicidad del caso lineal al no lineal.

TEOREMA 7.19. EXISTENCIA Y UNICIDAD DE LAS SOLUCIONES DE SISTEMAS NO LINEALES DE PRIMER ORDEN. *Supongamos que F satisface la acotación, la conti-*

nuidad y las condiciones de Lipschitz antes citadas en un conjunto S. Sea I el intervalo abierto $(a - c, a + c)$, donde $c =$ min $\{h, k/M\}$. Existe entonces una y sólo una función Y definida en I, siendo $Y(a) = B$, tal que $(x, Y(x)) \in S$ y

$$Y'(x) = F(x, Y(x)) \quad \textit{para cada } x \textit{ de } I.$$

Demostración. Puesto que la demostración es análoga a la del caso lineal indicamos solamente los principales pasos. Pongamos $Y_0(x) = B$ y definamos las funciones vectoriales $Y_1, Y_2, \ldots$ en I mediante la fórmula de recurrencia

$$(7.69) \qquad Y_{m+1}(x) = B + \int_a^x F[t, Y_m(t)]\, dt \quad \text{para} \quad m = 0, 1, 2, \ldots .$$

Para que la fórmula de recurrencia tenga sentido necesitamos saber si $(x, Y_m(x)) \in S$ para cada x de I. Esto se demuestra fácilmente por inducción en m. Cuando $m = 0$ tenemos $(x, Y_0(x)) = (x, B)$, que está en S. Supongamos entonces que $(x, Y_m(x)) \in S$ para un m y cada x de I. Utilizando (7.69) y (7.68) obtenemos

$$\| Y_{m+1}(x) - B \| \leq \left| \int_a^x \|F[t, Y_m(t)]\|\, dt \right| \leq M \left| \int_a^x dt \right| = M\,|x - a|.$$

Puesto que $|x - a| \leqslant c$ para x en I, esto implica que

$$\| Y_{m+1}(x) - B \| \leq Mc \leq k,$$

lo cual demuestra que $(x, Y_{m+1}(x)) \in S$ para cada x en I. Por consiguiente la fórmula de recurrencia tiene sentido para todo $m \geqslant 0$ y toda x de I.

La convergencia de la sucesión $\{Y_m(x)\}$ se establece ahora exactamente como en la sección 7.21. Escribamos

$$Y_k(x) = Y_0(x) + \sum_{m=0}^{k-1} \{Y_{m+1}(x) - Y_m(x)\}$$

y demostremos que $Y_k(x)$ tiende hacia un límite cuando $k \to \infty$ demostrando que la serie

$$\sum_{m=0}^{\infty} \| Y_{m+1}(x) - Y_m(x) \|$$

converge en I. Esto se deduce de la desigualdad

$$\| Y_{m+1}(x) - Y_m(x) \| \leq \frac{MA^m\,|x - a|^{m+1}}{(m + 1)!} \leq \frac{MA^m c^{m+1}}{(m + 1)!}$$

que se demuestra por inducción, usando la fórmula de recurrencia y la condición de Lipschitz. Definimos entonces la función límite Y por

$$Y(x) = \lim_{m \to \infty} Y_m(x)$$

para cada x de I y comprobamos que satisface la ecuación integral

$$Y(x) = B + \int_a^x F[t, Y(t)]\, dt,$$

exactamente como en el caso lineal. Esto demuestra la existencia de una solución. La unicidad puede entonces demostrarse por el mismo método utilizado para demostrar el teorema 7.17.

7.24 Ejercicios

1. Considerar el problema lineal de valor inicial

$$y' + y = 2e^x, \qquad \text{con } y = 1 \text{ cuando } x = 0.$$

a) Hallar la solución exacta Y de este problema.
b) Aplicar el método de aproximaciones sucesivas, partiendo de la función inicial $Y_0(x) = 1$.
Determinar $Y_n(x)$ explícitamente y demostrar que

$$\lim_{n \to \infty} Y_n(x) = Y(x)$$

para todo x real.

2. Aplicar el método de aproximaciones sucesivas al problema de valor inicial no lineal

$$y' = x + y^2, \qquad \text{con } y = 0 \text{ cuando } x = 0.$$

Tomar $Y_0(x) = 0$ como función inicial y calcular $Y_3(x)$.

3. Aplicar el método de aproximaciones sucesivas al problema de valor inicial no lineal

$$y' = 1 + xy^2, \qquad \text{con } y = 0 \text{ cuando } x = 0.$$

Tomar $Y_0(x) = 0$ como función inicial y calcular $Y_3(x)$.

4. Aplicar el método de aproximaciones sucesivas al problema de valor inicial no lineal

$$y' = x^2 + y^2, \qquad \text{con } y = 0 \text{ cuando } x = 0.$$

Partir de la «mala» función inicial $Y_0(x) = 1$, calcular $Y_3(x)$, y comparar con los resultados del ejemplo 1 de la sección 7.22.

5. Considerar el problema de valor inicial no lineal

$$y' = x^2 + y^2, \qquad \text{con } y = 1 \text{ cuando } x = 0.$$

a) Aplicar el método de aproximaciones sucesivas, partiendo de la función inicial $Y_0(x) = 1$, y calcular $Y_2(x)$.
b) Sea $R = [-1, 1] \times [-1, 1]$. Hallar el menor M tal que $|f(x, y)| \leq M$ en R. Hallar un intervalo $I = (-c, c)$ tal que la gráfica de toda función aproximante Y_n en I, esté en R.
c) Supóngase que la solución $y = Y(x)$ tiene un desarrollo en serie de potencias en un entorno del origen. Determinar los seis primeros términos no nulos de ese desarrollo y comparar el resultado con el del apartado a).

6. Considerar el problema de valor inicial

$$y' = 1 + y^2, \qquad \text{con } y = 0 \text{ cuando } x = 0.$$

a) Aplicar el método de aproximaciones sucesivas, partiendo de la función inicial $Y_0(x) = 0$, y calcular $Y_4(x)$.
b) Demostrar que toda función aproximante Y_n está definida en todo el eje real.
c) Utilizar el teorema 7.19 para demostrar que el problema de valor inicial tiene por lo menos una solución en cualquier intervalo de la forma $(-h, h)$.
d) Resolver la ecuación diferencial por separación de variables y demostrar así que existe una sola solución Y del problema de valor inicial en el intervalo $(-\pi/2, \pi/2)$ y ninguna en un intervalo más amplio. En este ejemplo, las aproximaciones sucesivas están definidas en todo el eje real, pero convergen hacia una función límite únicamente en el intervalo $(-\pi/2, \pi/2)$.

7. Busquemos dos funciones $y = Y(x)$ y $z = Z(x)$ que satisfagan simultáneamente el sistema de ecuaciones

$$y' = z, \qquad z' = x^3(y + z)$$

con las condiciones iniciales $y = 1$ y $z = 1/2$ cuando $x = 0$. Comenzar con las funciones iniciales $Y_0(x) = 1$, $Z_0(x) = 1/2$, y emplear el método de aproximaciones sucesivas para obtener las funciones aproximantes

$$Y_3(x) = 1 + \frac{x}{2} + \frac{3x^5}{40} + \frac{x^6}{60} + \frac{x^9}{192},$$

$$Z_3(x) = \frac{1}{2} + \frac{3x^4}{8} + \frac{x^5}{10} + \frac{3x^8}{64} + \frac{7x^9}{360} + \frac{x^{12}}{256}.$$

8. Considerar el sistema de ecuaciones

$$y' = 2x + z, \qquad z' = 3xy + x^2 z,$$

con valores iniciales $y = 2$ y $z = 0$ cuando $x = 0$. Partir de las funciones iniciales $Y_0(x) = 2$, $Z_0(x) = 0$, y con el método de aproximaciones sucesivas determinar $Y_3(x)$ y $Z_3(x)$.

9. Considerar el problema de valor inicial

$$y'' = x^2y' + x^4y, \qquad \text{con} \quad y = 5 \quad \text{e} \quad y' = 1 \text{ cuando } x = 0.$$

Cambiar este problema en otro equivalente que incluya un sistema de dos ecuaciones con dos funciones incógnitas $y = Y(x)$ y $z = Z(x)$, siendo $z = y'$. Aplicar entonces el método de aproximaciones sucesivas, partiendo de las funciones iniciales $Y_0(x) = 5$ y $Z_0(x) = 1$, y determinar $Y_3(x)$ y $Z_3(x)$.

10. Sea f una función definida en el rectángulo $R = [-1, 1] \times [-1, 1]$ del modo siguiente:

$$f(x, y) = \begin{cases} 0 & \text{si } x = 0, \\ 2y/x & \text{si } x \neq 0 \text{ e } |y| \leq x^2, \\ 2x & \text{si } x \neq 0 \text{ e } y > x^2, \\ -2x & \text{si } x \neq 0 \text{ e } y < -x^2. \end{cases}$$

a) Demostrar que $|f(x,y)| \leq 2$ para todo (x, y) de R.
b) Demostrar que f no satisface una condición de Lipschitz en R.
c) Para cada constante C que cumple $|C| \leq 1$, demostrar que $y = Cx^2$ es una solución del problema de valor inicial $y' = f(x, y)$, con $y = 0$ cuando $x = 0$. Demostrar que la gráfica de cada una de esas soluciones en $(-1, 1)$ está en R.
d) Aplicar el método de aproximaciones sucesivas al problema de valor inicial, partiendo de $Y_0(x) = 0$. Determinar $Y_n(x)$ y demostrar que las aproximaciones convergen hacia una solución del problema en el intervalo $(-1, 1)$.
e) Repetir la parte d), partiendo de $Y_0(x) = x$. Determinar $Y_n(x)$ y demostrar que las funciones aproximantes convergen hacia una solución diferente de cualquiera de las del apartado c).
f) Repetir la parte d), partiendo de $Y_0(x) = x^3$.
g) Repetir la parte d), partiendo de $Y_0(x) = x^{1/3}$.

* 7.25 Aproximaciones sucesivas y puntos fijos de operadores

La idea básica del método de aproximaciones sucesivas puede usarse no sólo para establecer teoremas de existencia para ecuaciones diferenciales sino también para otros muchos problemas del análisis. El resto de este capítulo reformula el citado método en una forma que amplía mucho el alcance de sus aplicaciones.

En la demostración del teorema 7.18 construimos una sucesión de funciones $\{Y_k\}$ de acuerdo con la fórmula de recurrencia

$$Y_{k+1}(x) = B + \int_a^x AY_k(t)\, dt.$$

El segundo miembro de esta fórmula puede considerarse como un operador T que convierte ciertas funciones Y en otras nuevas $T(Y)$ mediante la ecuación

$$T(Y) = B + \int_a^x AY(t)\,dt\,.$$

En la demostración del teorema 7.18 encontramos que la solución Y del problema de valor inicial $Y'(t) = AY(t)$, $Y(a) = B$, satisface la ecuación integral

$$Y = B + \int_a^x AY(t)\,dt\,.$$

Expresando ésta con el operador T resulta $Y = T(Y)$. Es decir, la solución Y permanece inalterada por el operador T. Una tal función Y se llama un *punto fijo* del operador T.

Numerosos e importantes problemas del análisis pueden formularse de modo que su solución dependa de la existencia de un punto fijo para algún operador. Por consiguiente merece la pena intentar descubrir propiedades de operadores que garanticen la existencia de un punto fijo. Volvamos ahora a un tratamiento sistemático de este problema.

* 7.26 Espacios lineales normados

Para formular el método de aproximaciones sucesivas en una forma más general conviene trabajar con espacios lineales. Sea S un espacio lineal cualquiera. Cuando hablamos de aproximar un elemento x de S mediante otro elemento y de S, consideramos la diferencia $x - y$, que llamamos el *error* de aproximación. Para medir el tamaño de ese error introducimos una norma en el espacio.

DEFINICIÓN DE NORMA. *Sea S un espacio lineal cualquiera. Una función real N definida en S se llama norma si tiene las propiedades siguientes:*

a) $N(x) \geq 0$ *para cada x de S.*
b) $N(cx) = |c|\, N(x)$ *para cada x de S y cada escalar c.*
c) $N(x + y) \leq N(x) + N(y)$ *para todo par x e y de S.*
d) $N(x) = 0$ *implica $x = O$.*

Un espacio lineal con una norma asignada se llama *espacio lineal normado.*

La norma de x a veces se indica con $\|x\|$ en lugar de $N(x)$. Con esta notación las propiedades fundamentales se expresan:

a) $\|x\| \geq 0$ para cada x de S.
b) $\|cx\| = |c|\, \|x\|$ para todo x de S y todo escalar c.

c) $\|x + y\| \leqslant \|x\| + \|y\|$ para todo par x e y de S.

d) $\|x\| = 0$ implica $x = O$.

Si x e y están en S, la norma $\|x - y\|$ se llama *distancia* de x a y.

Si el espacio S es euclídeo, tiene siempre una norma que es consecuencia del producto interior, es decir, $\|x\| = (x, x)^{1/2}$. Sin embargo, nosotros nos interesaremos por una norma especial que no resulta de un producto interior.

EJEMPLO. *La norma del máximo o* max *norma.* Sea $C(J)$ el espacio lineal de las funciones reales continuas en un intervalo cerrado y acotado J. Si $\varphi \in C(J)$, definimos

$$\|\varphi\| = \max_{x \in J} |\varphi(x)|,$$

donde el símbolo del segundo miembro significa el máximo valor absoluto de φ en J. El lector puede comprobar que esta norma goza de las cuatro propiedades fundamentales.

La norma del máximo no se deduce del producto interior. Para demostrarlo basta hacer ver que la norma del máximo deja de cumplir alguna propiedad que poseen todas las normas deducidas de productos interiores. Por ejemplo, si una norma se deduce de un producto interior, la «ley del paralelogramo»

$$\|x + y\|^2 + \|x - y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$$

es válida para todo par x e y de S. (Ver ejercicio 16 de la sección 1.13). La ley del paralelogramo no siempre queda satisfecha por la norma del máximo. Por ejemplo, sean x e y funciones en el intervalo $[0,1]$ dadas por

$$x(t) = t, \qquad y(t) = 1 - t.$$

Entonces tenemos $\|x\| = \|y\| = \|x + y\| = \|x - y\| = 1$, con lo que la ley del paralelogramo no se cumple.

* 7.27 Operadores de contracción

En esta sección consideramos el espacio lineal normado $C(J)$ de todas las funciones reales continuas en un intervalo cerrado y acotado J en el que $\|\varphi\|$ es la norma del maximo. Consideremos un operador

$$T: C(J) \rightarrow C(J)$$

cuyo dominio es $C(J)$ y cuyo recorrido es un subconjunto de $C(J)$. Esto es, si φ es continua en J, entonces $T(\varphi)$ también es continua en J. Las siguientes fórmulas dan algunos ejemplos sencillos de tales operadores. En cada caso φ es una función

arbitraria en $C(J)$ y $T(\varphi)(x)$ se define para cada x de J por la fórmula que se indica

$$T(\varphi)(x) = \lambda\varphi(x), \qquad \text{donde } \lambda \text{ es un número real fijo,}$$

$$T(\varphi)(x) = \int_c^x \varphi(t)\, dt, \qquad \text{donde } c \text{ es un punto fijo en } J,$$

$$T(\varphi)(x) = b + \int_c^x f[t, \varphi(t)]\, dt,$$

siendo b una constante y la función compuesta $f[t, \varphi(t)]$ continua en J.

Nos interesan ahora aquellos operadores T para los que la distancia $\|T(\varphi) - T(\psi)\|$ es menor que el producto de $\|\varphi - \psi\|$ por una constante $\alpha < 1$. Esos se llaman *operadores de contracción;* se definen del modo siguiente.

DEFINICIÓN DE UN OPERADOR DE CONTRACCIÓN. *Un operador T: $C(J)\to C(J)$ se llama de contracción si existe una constante α, que sea $0 \leqslant \alpha < 1$, tal que para todo par de funciones φ y ψ de $C(J)$ tenemos*

$$\|T(\varphi) - T(\psi)\| \leq \alpha \, \|\varphi - \psi\| . \tag{7.70}$$

La constante α se llama constante de contracción para T.

Nota: La desigualdad (7.70) es válida si y solamente si

$$|T(\varphi)(x) - T(\psi)(x)| \leq \alpha \, \|\varphi - \psi\| \qquad \text{para todo } x \text{ de } J.$$

EJEMPLO 1. Sea T el operador definido por $T(\varphi)(x) = \lambda\varphi(x)$, donde λ es constante. Puesto que

$$|T(\varphi)(x) - T(\psi)(x)| = |\lambda| \, |\varphi(x) - \psi(x)|,$$

tenemos $\|T(\varphi) - T(\psi)\| = |\lambda| \, \|\varphi - \psi\|$. Por consiguiente este operador es un operador de contracción si y sólo si $|\lambda| < 1$, en cuyo caso $|\lambda|$ puede usarse como constante de contracción.

EJEMPLO 2. Sea $T(\varphi)(x) = b + \int_c^x f[t, \varphi(t)]\, dt$, donde f satisface una condición de Lipschitz de la forma

$$|f(x, y) - f(x, z)| \leq K\, |y - z|$$

para todo x de J, y cualesquiera z e y reales; K es una constante positiva. Desig-

nemos con $L(J)$ la longitud del intervalo J. Si $KL(J) < 1$ podemos ver fácilmente que T es un operador de contracción con la constante de contracción $KL(J)$. En efecto, para toda x de J tenemos

$$|T(\varphi)(x) - T(\psi)(x)| = \left| \int_c^x \{f[t, \varphi(t)] - f[t, \psi(t)]\} \, dt \right| \leq K \left| \int_c^x |\varphi(t) - \psi(t)| \, dt \right|$$
$$\leq K \, \|\varphi - \psi\| \left| \int_c^x dt \right| \leq KL(J) \, \|\varphi - \psi\| .$$

Si $KL(J) < 1$, entonces T es un operador de contracción con constante de contracción $\alpha = KL(J)$.

* 7.28 Teorema del punto fijo para operadores de contracción

El teorema siguiente demuestra que todo operador de contracción tiene un punto fijo único.

TEOREMA 7.20. *Sea $T: C(J) \to C(J)$ un operador de contracción. Existe una función y sólo una φ en $C(J)$ tal que*

$$T(\varphi) = \varphi . \tag{7.71}$$

Demostración. Sea φ_0 cualquier función de $C(J)$ y definamos una sucesión de funciones $\{\varphi_n\}$ mediante la fórmula de recurrencia

$$\varphi_{n+1} = T(\varphi_n) \qquad \text{para} \quad n = 0, 1, 2, \ldots .$$

Obsérvese que $\varphi_{n+1} \in C(J)$ para cada n. Demostraremos que la sucesión $\{\varphi_n\}$ converge hacia una función límite φ de $C(J)$. El método es parecido al utilizado en la demostración del teorema 7.18. Escribimos cada φ_n como suma telescópica,

$$\varphi_n(x) = \varphi_0(x) + \sum_{k=0}^{n-1} \{\varphi_{k+1}(x) - \varphi_k(x)\} \tag{7.72}$$

y probamos la convergencia de $\{\varphi_n\}$ demostrando que la serie

$$\varphi_0(x) + \sum_{k=0}^{\infty} \{\varphi_{k+1}(x) - \varphi_k(x)\} \tag{7.73}$$

converge uniformemente en J. Luego demostramos que la suma de esa serie es el punto fijo que se cita.

La convergencia uniforme de la serie se establecerá comparándola con la serie geométrica convergente

$$M\sum_{k=0}^{\infty}\alpha^k,$$

donde $M = \|\varphi_0\| + \|\varphi_1\|$, y α es la constante de contracción para T. La comparación la proporciona la desigualdad

$$|\varphi_{k+1}(x) - \varphi_k(x)| \le M\alpha^k \tag{7.74}$$

que es válida para toda x de J y todo $k \geqslant 1$. Para demostrar (7.74) notemos que

$$|\varphi_{k+1}(x) - \varphi_k(x)| = |T(\varphi_k)(x) - T(\varphi_{k-1})(x)| \le \alpha\,\|\varphi_k - \varphi_{k-1}\|\,.$$

Por consiguiente la desigualdad (7.74) quedará probada si demostramos que

$$\|\varphi_k - \varphi_{k-1}\| \le M\alpha^{k-1} \tag{7.75}$$

para todo $k \geqslant 1$. Demostremos ahora (7.75) por inducción. Para $k = 1$ tenemos

$$\|\varphi_1 - \varphi_0\| \le \|\varphi_1\| + \|\varphi_0\| = M,$$

que es la misma que (7.75). Para demostrar que (7.75) es válida para $k + 1$ si lo es para k observemos que

$$|\varphi_{k+1}(x) - \varphi_k(x)| = |T(\varphi_k)(x) - T(\varphi_{k-1})(x)| \le \alpha\,\|\varphi_k - \varphi_{k-1}\| \le M\alpha^k.$$

Puesto que ésta es válida para cada x de J también debe ser

$$\|\varphi_{k+1} - \varphi_k\| \le M\alpha^k.$$

Esto demuestra (7.75) por inducción. Por consiguiente la serie (7.73) converge uniformemente en J. Si designamos por $\varphi(x)$ su suma tenemos

$$\varphi(x) = \lim_{n\to\infty}\varphi_n(x) = \varphi_0(x) + \sum_{k=0}^{\infty}\{\varphi_{k+1}(x) - \varphi_k(x)\}\,. \tag{7.76}$$

La función φ es continua en J porque es la suma de una serie uniformemente

convergente de funciones continuas. Para demostrar que φ es un punto fijo de T comparamos $T(\varphi)$ con $\varphi_{n+1} = T(\varphi_n)$. Utilizando la propiedad de contracción de T tenemos

$$(7.77) \qquad |T(\varphi)(x) - \varphi_{n+1}(x)| = |T(\varphi)(x) - T(\varphi_n)(x)| \leq \alpha\, |\varphi(x) - \varphi_n(x)|\,.$$

Pero de (7.72) y (7.76) encontramos

$$|\varphi(x) - \varphi_n(x)| = \left|\sum_{k=n}^{\infty}\{\varphi_{k+1}(x) - \varphi_k(x)\}\right| \leq \sum_{k=n}^{\infty}|\varphi_{k+1}(x) - \varphi_k(x)| \leq M\sum_{k=n}^{\infty}\alpha^k,$$

donde hemos utilizado (7.74) en el último paso. Por consiguiente (7.77) implica

$$|T(\varphi)(x) - \varphi_{n+1}(x)| \leq M\sum_{k=n}^{\infty}\alpha^{k+1}.$$

Cuando $n \to \infty$ la serie del segundo miembro tiende a 0, así que $\varphi_{n+1}(x) \to T(\varphi)(x)$. Pero ya que $\varphi_{n+1}(x) \to \varphi(x)$ cuando $n \to \infty$, esto demuestra que $\varphi(x) = T(\varphi)(x)$ para cada x de J. Por consiguiente $\varphi = T(\varphi)$, así que φ es un punto fijo.

Por último demostramos que este punto fijo φ es único. Sea ψ otra función de $C(J)$ tal que $T(\psi) = \psi$. Tenemos entonces

$$\|\varphi - \psi\| = \|T(\varphi) - T(\psi)\| \leq \alpha\, \|\varphi - \psi\|\,.$$

Esto nos da $(1 - \alpha)\|\varphi - \psi\| \leqslant 0$. Puesto que $\alpha < 1$ podemos dividir por $1 - \alpha$ obteniendo la desigualdad $\|\varphi - \psi\| \leqslant 0$. Pero como también tenemos $\|\varphi - \psi\| \geqslant 0$ ello significa que $\|\varphi - \psi\| = 0$, y por tanto $\varphi - \psi = 0$. La demostración del teorema del punto fijo ha quedado concluida.

* 7.29 Aplicaciones del teorema del punto fijo

Para hacer ver el amplio alcance de las aplicaciones del teorema del punto fijo lo vamos a utilizar para demostrar dos teoremas importantes. El primero da una condición suficiente para que una ecuación de la forma $f(x, y) = 0$ defina y como función de x.

TEOREMA 7.21. UN TEOREMA DE EXISTENCIA DE LA FUNCIÓN IMPLÍCITA. *Sea f una función definida en una banda R de la forma*

$$R = \{(x, y) \mid a \leq x \leq b\,,\ -\infty < y < +\infty\}\,.$$

Supongamos que existe la derivada parcial $D_2f(x, y)$(*) y que satisface una desigualdad de la forma

$$0 < m \leq D_2f(x, y) \leq M \tag{7.78}$$

para todo punto (x, y) de R, donde m y M son constantes tales que $m \leqslant M$. Supongamos también que para cada función φ continua en $[a, b]$ la función compuesta $g(x) = f[x, \varphi(x)]$ es continua en $[a, b]$. Entonces existe una y sólo una función $y = Y(x)$, continua en $[a, b]$ tal que

$$f[x, Y(x)] = 0 \tag{7.79}$$

para todo x en $[a, b]$.

Observación: Expresamos este resultado diciendo que la ecuación $f(x, y) = 0$ sirve para definir y implícitamente como función de x en $[a, b]$.

Demostración. Designemos con C el espacio lineal de las funciones continuas en $[a, b]$, y definamos un operador T: $C \to C$ mediante la ecuación

$$T(\varphi)(x) = \varphi(x) - \frac{1}{M} f[x, \varphi(x)]$$

para cada x de $[a, b]$. La constante positiva M es la de (7.78). La función $T(\varphi) \in C$ siempre que $\varphi \in C$. Demostraremos que T es un operador de contracción. Una vez sabido esto se deducirá que T tiene un punto fijo único Y en C. Para esta función Y tendremos $Y = T(Y)$ lo cual significa

$$Y(x) = Y(x) - \frac{1}{M} f[x, Y(x)]$$

para cada x de $[a, b]$. Esto nos dará (7.79), como queríamos.

Para demostrar que T es un operador de contracción consideremos la diferencia

$$T(\varphi)(x) - T(\psi)(x) = \varphi(x) - \psi(x) - \frac{f[x, \varphi(x)] - f[x, \psi(x)]}{M}. \tag{7.80}$$

Según el teorema del valor medio para la derivación tenemos

$$f[x, \varphi(x)] - f[x, \psi(x)] = D_2f[x, z(x)][\varphi(x) - \psi(x)],$$

(*) $D_2f(x, y)$ es la derivada de $f(x, y)$ respecto a y, manteniendo x fijo.

donde $z(x)$ está situada entre $\varphi(x)$ y $\psi(x)$. Por consiguiente (7.80) nos da

$$(7.81) \qquad T(\varphi)(x) - T(\psi)(x) = [\varphi(x) - \psi(x)]\left(1 - \frac{D_2 f[x, z(x)]}{M}\right).$$

La hipótesis (7.78) implica que

$$0 \leq 1 - \frac{D_2 f[x, z(x)]}{M} \leq 1 - \frac{m}{M}.$$

En consecuencia (7.81) nos da la desigualdad

$$(7.82) \qquad |T(\varphi)(x) - T(\psi)(x)| \leq |\varphi(x) - \psi(x)| \left(1 - \frac{m}{M}\right) \leq \alpha \, \|\varphi - \psi\|,$$

donde $\alpha = 1 - m/M$. Puesto que $0 < m \leqslant M$, tenemos $0 \leqslant \alpha < 1$. La desigualdad (7.82) es válida para todo x de $[a, b]$. Luego T es un operador de contracción. Esto completa la demostración.

La otra aplicación del teorema del punto fijo establece un teorema de existencia para la ecuación integral

$$(7.83) \qquad \varphi(x) = \psi(x) + \lambda \int_a^b K(x, t)\varphi(t)\, dt.$$

Aquí ψ es una función dada, continua en $[a, b]$, λ es una constante dada, y K una función dada definida y acotada en el cuadrado

$$S = \{(x, y) \mid a \leq x \leq b, a \leq y \leq b\}.$$

La función K se llama el *núcleo* de la ecuación integral. Sea C el espacio lineal de las funciones continuas en $[a, b]$. Supongamos que el núcleo K es tal que el operador T dado por

$$T(\varphi)(x) = \psi(x) + \lambda \int_a^b K(x, t)\varphi(t)\, dt$$

aplica C en C. Es decir, supongamos que $T(\varphi) \in C$ siempre que $\varphi \in C$. Una solución de la ecuación integral es cualquier función φ de C que satisfaga (7.83).

TEOREMA 7.22. TEOREMA DE EXISTENCIA PARA ECUACIONES INTEGRALES. *Si, en las condiciones impuestas, tenemos*

$$(7.84) \qquad |K(x, y)| \leq M$$

para todo punto (x, y) de S, siendo $M > 0$, entonces para cada λ tal que

$$(7.85) \qquad |\lambda| < \frac{1}{M(b - a)}$$

existe una y sólo una función φ en C que satisface la ecuación integral (7.83).

Demostración. Demostraremos que T es un operador de contracción. Tomemos en C dos funciones cualesquiera φ_1 y φ_2 y consideremos la diferencia

$$T(\varphi_1)(x) - T(\varphi_2)(x) = \lambda \int_a^b K(x, t)[\varphi_1(t) - \varphi_2(t)]\, dt .$$

Utilizando la desigualdad (7.48) podemos escribir

$$|T(\varphi_1)(x) - T(\varphi_2)(x)| \leq |\lambda|\, M(b - a)\, \|\varphi_1 - \varphi_2\| = \alpha\, \|\varphi_1 - \varphi_2\| ,$$

siendo $\alpha = |\lambda|\, M(b - a)$. Por (7.85) tenemos $0 \leqslant \alpha < 1$, así que T es un operador de contracción con α como constante de contracción. Por consiguiente T tiene un punto fijo φ único en C. Esta función φ satisface (7.83).

PARTE 2

Análisis no lineal

8

CALCULO DIFERENCIAL EN CAMPOS ESCALARES Y VECTORIALES

8.1 Funciones de $\mathbf{R}^n$ en $\mathbf{R}^m$. Campos escalares y vectoriales

La parte 1 de este volumen se dedica principalmente a las transformaciones lineales

$$T: V \to W$$

de un espacio lineal V en otro espacio lineal W. En la parte 2 prescindimos de que T sea lineal pero restringimos los espacios V y W a ser de dimensión finita. Concretamente, consideramos funciones con el dominio en el espacio n-dimensional $\mathbf{R}^n$ y con el recorrido en el espacio m-dimensional $\mathbf{R}^m$.

Cuando los dos números m y n son iguales a 1, una tal función se llama función real de una variable real. Cuando $n = 1$ y $m > 1$ se llama función vectorial de una variable real. En el Volumen I se estudiaron ampliamente ejemplos de tales funciones.

En este capítulo suponemos $n > 1$ y $m \geqslant 1$. Cuando $m = 1$ la función se llama función real de una variable vectorial o, más brevemente, un *campo escalar*. Cuando $m > 1$ se llama función vectorial de una variable vectorial, o simplemente *campo vectorial.*

Este capítulo extiende los conceptos de límite, continuidad y derivada, a los campos escalares y vectoriales. Los Capítulos 10 y 11 extienden el concepto de integral.

Notación: Los escalares se designarán con tipo de letra corriente, y los vectores con tipo de letra negrita. Si f es un campo escalar definido en un punto $\boldsymbol{x} = (x_1, \ldots, x_n)$ en $\mathbf{R}^n$, la notación $f(\boldsymbol{x})$ y $f(x_1, \ldots, x_n)$ se usan indistintamente para designar el valor de f en un cierto punto. Si $\boldsymbol{f}$ es un campo vectorial ponemos $\boldsymbol{f}(\boldsymbol{x})$ o $\boldsymbol{f}(x_1, \ldots, x_n)$ para el valor de la función en $\boldsymbol{x}$. Utilizaremos el producto interior

$$x \cdot y = \sum_{k=1}^{n} x_k y_k$$

y la correspondiente norma $\|x\| = (x \cdot x)^{1/2}$, donde $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$. Los puntos de $\mathbf{R}^2$ ordinariamente se indican por (x, y) en lugar de (x_1, x_2); y los de $\mathbf{R}^3$ por (x, y, z) en lugar de (x_1, x_2, x_3).

Los campos escalares y vectoriales definidos en subconjunto de $\mathbf{R}^2$ y $\mathbf{R}^3$ se presentan con frecuencia en las aplicaciones de la Matemática a las Ciencias naturales y a la Ingeniería. Por ejemplo, si a cada punto x de la atmósfera le asignamos un número real $f(x)$ que represente la temperatura en x, la función f así definida es un campo escalar. Si asignamos un vector que represente la velocidad del viento en aquel punto, obtenemos un ejemplo de campo vectorial.

En problemas de Física que tratan con campos escalares o vectoriales es importante saber cómo varía el campo al pasar de un punto a otro. En el caso uni-dimensional la derivada es el instrumento matemático utilizado para estudiar tales cambios. La teoría de la derivada en el caso uni-dimensional maneja funciones definidas en intervalos abiertos. Para extender la teoría a $\mathbf{R}^n$ consideramos generalizaciones de intervalos abiertos llamados *conjuntos abiertos.*

8.2 Bolas abiertas y conjuntos abiertos

Sean a un punto dado en $\mathbf{R}^n$ y r un número positivo dado. El conjunto de todos los puntos x de $\mathbf{R}^n$ tales que

$$\|x - a\| < r$$

se llama *n-bola abierta* de radio r y centro a. Designamos este conjunto con $B(a;r)$.

La bola $B(a;r)$ está constituida por los puntos cuya distancia a a es menor que r. En $\mathbf{R}^1$ es simplemente un intervalo abierto con punto medio en a. En $\mathbf{R}^2$ es un disco circular, y en $\mathbf{R}^3$ es un sólido esférico con centro en a y radio r.

DEFINICIÓN DE PUNTO INTERIOR. *Sea S un subconjunto de $\mathbf{R}^n$, y supongamos que $a \in S$. Se dice que a es punto interior de S si existe una n-bola abierta con centro en a, cuyos puntos pertenecen todos a S.*

Es decir, todo punto a interior de S puede rodearse por una n-bola $B(a)$ tal que $B(a) \subseteq S$. El conjunto de todos los puntos interiores de S se llaman el *interior* de S y se designa con in S. Un conjunto abierto que contenga un punto a se llama a veces *entorno* de a.

DEFINICIÓN DE CONJUNTO ABIERTO. *Un conjunto S de $\mathbf{R}^n$ se llama abierto si todos sus puntos son interiores. Es decir, S es abierto si y sólo si $S = \operatorname{int} S$.*

EJEMPLOS. En $\mathbf{R}^1$ el tipo más sencillo de conjunto abierto es un intervalo abierto. La reunión de dos o más intervalos abiertos es también abierto. Un intervalo cerrado $[a, b]$ no es un conjunto abierto porque ninguno de los extremos del intervalo puede encerrarse en una 1-bola situada enteramente en el intervalo dado.

La 2-bola $S = B(O;1)$ dibujada en la figura 8.1 es un ejemplo de conjunto abierto en $\mathbf{R}^2$. Todo punto $\boldsymbol{a}$ de S es el centro de un disco situado por entero en S. Para algunos puntos el radio de este disco es muy pequeño.

Algunos conjuntos abiertos en $\mathbf{R}^2$ pueden construirse tomando el producto cartesiano de conjuntos abiertos en $\mathbf{R}^1$. Si A_1 y A_2 son subconjuntos de $\mathbf{R}^1$, su producto cartesiano $A_1 \times A_2$ es el conjunto definido en $\mathbf{R}^2$ por

$$A_1 \times A_2 = \{(a_1, a_2) \mid a_1 \in A_1 \quad \text{and} \quad a_2 \in A_2\}.$$

En la figura 8.2 se ha dibujado un ejemplo. Los conjuntos A_1 y A_2 son intervalos, y $A_1 \times A_2$ es un rectángulo.

Si A_1 y A_2 son subconjuntos abiertos de $\mathbf{R}^1$, entonces $A_1 \times A_2$ será un subconjunto abierto de $\mathbf{R}^2$. Para probar esto, elijamos un punto cualquiera $\boldsymbol{a} = (a_1, a_2)$

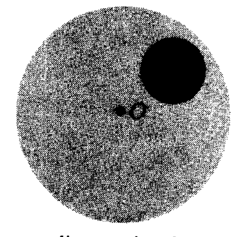


FIGURA 8.1 *El disco $B(O;\,1)$ es un conjunto abierto en $\mathbf{R}^2$*

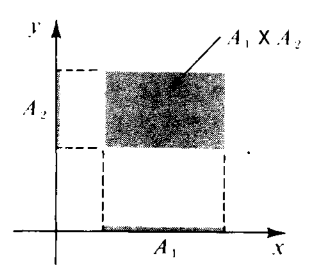


FIGURA 8.2 *El producto cartesiano de dos intervalos abiertos es un rectángulo abierto.*

en $A_1 \times A_2$. Tenemos que demostrar que $\boldsymbol{a}$ es un punto interior de $A_1 \times A_2$. Puesto que A_1 y A_2 son abiertos en $\mathbf{R}^1$ existe una 1-bola $B(a_1;\, r_1)$ en A_1 y una 1-bola $B(a_2;\, r_2)$ en A_2. Sea $r = \min\{r_1, r_2\}$. Podemos fácilmente demostrar que la 2-bola $(B(\boldsymbol{a}; r) \subseteq A_1 \times A_2$. En efecto, si $\boldsymbol{x} = (x_1, x_2)$ es un punto cualquiera de $B(\boldsymbol{a};\, r)$ entonces $\|\boldsymbol{x} - \boldsymbol{a}\| < r$, así que $|x_1 - a_1| < r_1$ y $|x_2 - a_2| < r_2$. Luego $x_1 \in B(a_1;\, r_1)$ y $x_2 \in B(a; r_2)$. Por consiguiente $x_1 \in A_1$ y $x_2 \in A_2$, así que $(x_1, x_2) \in A_1 \times A_2$. Esto demuestra que todo punto de $B(\boldsymbol{a};\, r)$ está en $A_1 \times A_2$. Por lo tanto todo punto de $A_1 \times A_2$ es un punto interior, así que $A_1 \times A_2$ es abierto.

El lector debería comprobar que un subconjunto abierto de $\mathbf{R}^1$ ya no es un conjunto abierto cuando se considera como subconjunto de $\mathbf{R}^2$, porque un subconjunto de $\mathbf{R}^1$ no puede contener una 2-bola.

DEFINICIONES DE EXTERIOR Y FRONTERA. *Un punto x se llama exterior al conjunto S de $\mathbf{R}^n$ si existe una n-bola $B(x)$ que no contiene puntos de S. El conjunto de todos los puntos de $\mathbf{R}^n$ exteriores a S se llama el exterior de S y se designa con* ext S. *Un punto que no es interior ni exterior a S se llama punto frontera de S. El conjunto de todos los puntos frontera de S es la frontera de S y se designa con ∂S.*

Estos conceptos se representan en la figura 8.1. El exterior de S es el conjunto de todos los x tales que $\|x\| > 1$. La frontera de S la constituyen todos los x con $\|x\| = 1$.

8.3 Ejercicios

1. Sean f un campo escalar definido en un conjunto S y c un número real dado. El conjunto de todos los puntos x de S tales que $f(x) = c$ se llama *conjunto de nivel* de f. (En un capítulo posterior se discutirán problemas geométricos y físicos en los que se presentan conjuntos de nivel.) Para cada uno de los campos escalares siguientes, S es todo el espacio $\mathbf{R}^n$. Hacer un dibujo para describir los conjuntos de nivel correspondientes a los valores dados de c.
   - a) $f(x, y) = x^2 + y^2$, $c = 0, 1, 4, 9$.
   - b) $f(x, y) = e^{xy}$, $c = e^{-2}, e^{-1}, 1, e, e^2, e^3$.
   - c) $f(x, y) = \cos(x + y)$, $c = -1, 0, \frac{1}{2}, \frac{1}{2}\sqrt{2}, 1$.
   - d) $f(x, y, z) = x + y + z$, $c = -1, 0, 1$.
   - e) $f(x, y, z) = x^2 + 2y^2 + 3z^2$, $c = 0, 6, 12$.
   - f) $f(x, y, z) = \operatorname{sen}(x^2 + y^2 + z^2)$, $c = -1, -\frac{1}{2}, 0, \frac{1}{2}\sqrt{2}, 1$.
2. En cada uno de los casos siguientes, sea S el conjunto de todos los puntos (x, y) del plano que satisfacen las desigualdades dadas. Hacer un gráfico mostrando el conjunto S y explicar, geométricamente, si S es o no abierto. Indicar la frontera de S en el gráfico.
   - a) $x^2 + y^2 < 1$.
   - b) $3x^2 + 2y^2 < 6$.
   - c) $|x| < 1$ y $|y| < 1$.
   - d) $x \geq 0$ y $y > 0$.
   - e) $|x| \leq 1$ y $|y| \leq 1$.
   - f) $x > 0$ y $y < 0$.
   - g) $xy < 1$.
   - h) $1 \leq x \leq 2$ y $3 < y < 4$.
   - i) $1 < x < 2$ y $y > 0$.
   - j) $x \geq y$.
   - k) $x > y$.
   - l) $y > x^2$ y $|x| < 2$.
   - m) $(x^2 + y^2 - 1)(4 - x^2 - y^2) > 0$.
   - n) $(2x - x^2 - y^2)(x^2 + y^2 - x) > 0$.
3. En cada uno de los siguientes casos, sea S el conjunto de todos los puntos (x, y, z), en el espacio tri-dimensional, que satisfacen las desigualdades dadas y determinar si S es o no abierto.
   - a) $z^2 - x^2 - y^2 - 1 > 0$.
   - b) $|x| < 1$, $|y| < 1$, y $|z| < 1$.
   - c) $x + y + z < 1$.
   - d) $|x| \leq 1$, $|y| < 1$, y $|z| < 1$.
   - e) $x + y + z < 1$ y $x > 0, y > 0, z > 0$.
   - f) $x^2 + 4y^2 + 4z^2 - 2x + 16y + 40z + 113 < 0$.

4. a) Si A es un conjunto abierto en un espacio de dimensión n y si $x \in A$, demostrar que el conjunto $A - \{x\}$, obtenido suprimiendo de A el punto x, es abierto.
   b) Si A es un intervalo abierto en la recta real y B un subintervalo de A cerrado, demostrar que $A - B$ es abierto. (*)
   c) Si A y B son intervalos abiertos en la recta real, demostrar que $A \cap B$ y $A \cup B$ son abiertos.
   d) Si A es un intervalo cerrado en la recta real, demostrar que su complemento (respecto a la recta real completa) es abierto.
5. Demostrar las siguientes propiedades de los conjuntos abiertos en $\mathbf{R}^n$:
   a) El conjunto vacío $\varnothing$ es abierto.
   b) $\mathbf{R}^n$ es abierto.
   c) La reunión de una colección cualquiera de conjuntos abiertos es un conjunto abierto.
   d) La intersección de una colección finita de conjuntos abiertos es un conjunto abierto.
   e) Dar un ejemplo para demostrar que la intersección de una colección infinita de conjuntos abiertos no es necesariamente abierta.

Conjuntos cerrados. Un conjunto S en $\mathbf{R}^n$ se llama *cerrado* si su complemento $\mathbf{R}^n - S$ es abierto. Los tres ejercicios siguientes se refieren a las propiedades de los conjuntos cerrados.

6. En cada uno de los siguientes casos, sea S el conjunto de todos los puntos (x, y) de $\mathbf{R}^2$ que satisfacen las condiciones dadas. Hacer un gráfico mostrando el conjunto S y dar una razón geométrica para explicar si S es abierto, cerrado, cerrado y abierto a la vez, o ni cerrado ni abierto.
   a) $x^2 + y^2 \geq 0$.
   b) $x^2 + y^2 < 0$.
   c) $x^2 + y^2 \leq 1$.
   d) $1 < x^2 + y^2 < 2$.
   e) $1 \leq x^2 + y^2 \leq 2$.
   f) $1 < x^2 + y^2 \leq 2$.
   g) $1 \leq x \leq 2, 3 \leq y \leq 4$.
   h) $1 \leq x \leq 2, 3 \leq y < 4$.
   i) $y = x^2$.
   j) $y \geq x^2$.
   k) $y \geq x^2$ y $|x| < 2$.
   l) $y \geq x^2$ y $|x| \leq 2$.
7. a) Si A es un conjunto cerrado en el espacio n-dimensional y x es un punto no perteneciente a A, demostrar que $A \cup \{x\}$ es también cerrado.
   b) Demostrar que un intervalo cerrado $[a, b]$ en la recta real es un conjunto cerrado.
   c) Si A y B son intervalos cerrados en la recta real, demostrar que $A \cap B$ y $A \cup B$ son cerrados.
8. Demostrar las propiedades siguientes de los conjuntos cerrados de $\mathbf{R}^n$. Pueden utilizarse los resultados del ejercicio 5.
   a) El conjunto vacío $\varnothing$ es cerrado.
   b) $\mathbf{R}^n$ es cerrado.
   c) La intersección de una colección cualquiera de conjuntos cerrados es cerrada.
   d) La reunión de un número finito de conjuntos cerrados es cerrada.
   e) Dar un ejemplo para probar que la reunión de una colección infinita de conjuntos cerrados no es necesariamente cerrada.
9. Sea S un subconjunto de $\mathbf{R}^n$.
   a) Demostrar que $\operatorname{int} S$ y $\operatorname{ext} S$ son ambos conjuntos abiertos.
   b) Demostrar que $\mathbf{R}^n = (\operatorname{int} S) \cup (\operatorname{ext} S) \cup \partial S$, es una reunión de conjuntos disjuntos, y utilizar esto para deducir que la frontera ∂S siempre es un conjunto cerrado.

(*) Si A y B son dos conjuntos, la diferencia $A - B$ (llamada *complemento de B respecto de A*) es el conjunto de todos los elementos de A que no pertenecen a B.

10. Dado un conjunto S de $\mathbf{R}^n$ y un punto $\boldsymbol{x}$ con la propiedad de que toda bola $B(x)$ contiene puntos interiores y exteriores a S simultáneamente. Demostrar que $\boldsymbol{x}$ es un punto frontera de S. ¿Es cierta la proposición inversa? Esto es, ¿todo punto frontera de S tiene necesariamente esta propiedad?
11. Sea S un subconjunto de $\mathbf{R}^n$. Demostrar que $\text{ext}\, S = \text{int}\,(\mathbf{R}^n - S)$.
12. Demostrar que un conjunto S de $\mathbf{R}^n$ es cerrado si y sólo si $S = (\text{int}\, S) \cup \partial S$.

8.4 Límites y continuidad

Los conceptos de límite y continuidad se pueden extender con facilidad a los campos escalares y vectoriales. Formularemos las definiciones para campos vectoriales; las mismas se extienden a los campos escalares.

Consideremos una función $\boldsymbol{f}\colon S \to \mathbf{R}^m$, donde S es un subconjunto de $\mathbf{R}^n$. Si $\boldsymbol{a} \in \mathbf{R}^n$ y $\boldsymbol{b} \in \mathbf{R}^m$ escribimos

(8.1) $$\lim_{\boldsymbol{x}\to\boldsymbol{a}} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{b} \qquad (\text{o, } \boldsymbol{f}(\boldsymbol{x}) \to \boldsymbol{b} \text{ cuando } \boldsymbol{x} \to \boldsymbol{a})$$

para significar que

(8.2) $$\lim_{\|\boldsymbol{x}-\boldsymbol{a}\|\to 0} \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{b}\| = 0.$$

El símbolo de límite de la igualdad (8.2) es el límite corriente del Cálculo elemental. En esta definición no se exige que $\boldsymbol{f}$ esté definida en el mismo punto $\boldsymbol{a}$.

Si escribimos $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{a}$, la igualdad (8.2) se convierte en

$$\lim_{\|\boldsymbol{h}\|\to 0} \|\boldsymbol{f}(\boldsymbol{a} + \boldsymbol{h}) - \boldsymbol{b}\| = 0.$$

Para los puntos de $\mathbf{R}^2$ escribimos (x, y) para $\boldsymbol{x}$ y (a, b) para $\boldsymbol{a}$ y expresamos la relación de límite (8.1) como sigue:

$$\lim_{(x,y)\to(a,b)} \boldsymbol{f}(x, y) = \boldsymbol{b}.$$

Para los puntos de $\mathbf{R}^3$ ponemos $\boldsymbol{x} = (x, y, z)$ y $\boldsymbol{a} = (a, b, c)$ y escribimos

$$\lim_{(x,y,z)\to(a,b,c)} \boldsymbol{f}(x, y, z) = \boldsymbol{b}.$$

Una función $\boldsymbol{f}$ se llama *continua* en $\boldsymbol{a}$ si $\boldsymbol{f}$ está definida en $\boldsymbol{a}$ y si

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{a}).$$

Decimos que $\boldsymbol{f}$ es *continua en un conjunto* S si $\boldsymbol{f}$ es continua en cada punto de S.

Puesto que esas definiciones son extensiones directas del caso uni-dimensional, no debe sorprender que muchas de las propiedades de los límites y de la continuidad pueden también extenderse. Por ejemplo, los teoremas relativos a los límites y a la continuidad de sumas, productos y cocientes son también ciertos para campos escalares. Para campos vectoriales, los cocientes no están definidos pero tenemos el teorema siguiente relativo a sumas, multiplicación por escalares, productos interiores y normas.

TEOREMA 8.1. *Si* $\lim_{x \to a} f(x) = b$ *y* $\lim_{x \to a} g(x) = c$, *tenemos también:*

a) $\lim_{x \to a} [f(x) + g(x)] = b + c$.

b) $\lim_{x \to a} \lambda f(x) = \lambda b$ *para todo escalar* λ.

c) $\lim_{x \to a} f(x) \cdot g(x) = b \cdot c$.

d) $\lim_{x \to a} \|f(x)\| = \|b\|$.

Demostración. Sólo vamos a demostrar las partes c) y d); las demostraciones de a) y b) se dejan como ejercicios para el lector.

Para demostrar c) escribimos

$$f(x) \cdot g(x) - b \cdot c = [f(x) - b] \cdot [g(x) - c] + b \cdot [g(x) - c] + c \cdot [f(x) - b].$$

Apliquemos ahora la desigualdad triangular y la desigualdad de Cauchy-Schwarz obteniendo

$$0 \leq \|f(x) \cdot g(x) - b \cdot c\| \leq \|f(x) - b\| \, \|g(x) - c\| + \|b\| \, \|g(x) - c\| + \|c\| \, \|f(x) - b\|.$$

Puesto que $\|f(x) - b\| \to 0$ y $\|g(x) - c\| \to 0$ cuando $x \to a$, esto prueba que $\|f(x) \cdot g(x) - b \cdot c\| \to 0$ cuando $x \to a$, lo cual demuestra c).

Tomando $f(x) = g(x)$ en el apartado c) encontramos

$$\lim_{x \to a} \|f(x)\|^2 = \|b\|^2,$$

de la que obtenemos d).

EJEMPLO 1. *Continuidad y componentes de un campo vectorial.* Si un campo vectorial f tiene los valores en $\mathbf{R}^m$, cada uno de los valores $f(x)$ tiene m componentes y podemos escribir

$$f(x) = (f_1(x), \ldots, f_m(x)).$$

Los m campos escalares $f_1, \ldots, f_m$ se llaman *componentes* del campo vectorial $\boldsymbol{f}$. Demostraremos que $\boldsymbol{f}$ es continuo en un punto si, y sólo si, cada componente f_k es continuo en dicho punto.

Designemos con $\boldsymbol{e}_k$ el k-ésimo vector coordenado unidad (todos los componentes de $\boldsymbol{e}_k$ son 0 excepto el k-ésimo, que es igual a 1). Entonces $f_k(\boldsymbol{x})$ viene dado por el producto interior

$$f_k(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{e}_k .$$

Por consiguiente, el apartado c) del teorema 8.1 demuestra que cada punto de continuidad de $\boldsymbol{f}$ lo es también de f_k. Además, puesto que tenemos

$$\boldsymbol{f}(\boldsymbol{x}) = \sum_{k=1}^{m} f_k(\boldsymbol{x})\boldsymbol{e}_k ,$$

la aplicación reiterada de los apartados a) y b) del teorema 8.1 demuestra que un punto de continuidad de todos los m componentes $f_1, \ldots, f_m$ es también punto de continuidad de $\boldsymbol{f}$.

EJEMPLO 2. *Continuidad de la función identidad.* La función identidad, $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}$, es continua en todo $\mathbf{R}^n$. Por tanto sus componentes también son continuos en todo $\mathbf{R}^n$. Esos son n campos escalares dados por

$$f_1(\boldsymbol{x}) = x_1, f_2(\boldsymbol{x}) = x_2, \ldots, f_n(\boldsymbol{x}) = x_n .$$

EJEMPLO 3. *Continuidad de las transformaciones lineales.* Sea $\boldsymbol{f}: \mathbf{R}^n \to \mathbf{R}^m$ una transformación lineal. Demostraremos que $\boldsymbol{f}$ es continua en cada punto $\boldsymbol{a}$ de $\mathbf{R}^n$. En virtud de la linealidad tenemos

$$\boldsymbol{f}(\boldsymbol{a} + \boldsymbol{h}) = \boldsymbol{f}(\boldsymbol{a}) + \boldsymbol{f}(\boldsymbol{h}).$$

Por consiguiente, basta demostrar que $\boldsymbol{f}(\boldsymbol{h}) \to \boldsymbol{O}$ cuando $\boldsymbol{h} \to \boldsymbol{O}$. Poniendo $\boldsymbol{h}$ en función de sus componentes tenemos $\boldsymbol{h} = h_1\boldsymbol{e}_1 + \cdots + h_n\boldsymbol{e}_n$. Usando la linealidad, tenemos $\boldsymbol{f}(\boldsymbol{h}) = h_1\boldsymbol{f}(\boldsymbol{e}_1) + \cdots + h_n\boldsymbol{f}(\boldsymbol{e}_n)$. Esto demuestra que $\boldsymbol{f}(\boldsymbol{h}) \to \boldsymbol{O}$ cuando $\boldsymbol{h} \to \boldsymbol{O}$.

EJEMPLO 4. *Continuidad de los polinomios de n variables.* Un campo escalar P definido en $\mathbf{R}^n$ por una fórmula de la forma

$$P(\boldsymbol{x}) = \sum_{k_1=0}^{p_1} \cdots \sum_{k_n=0}^{p_n} c_{k_1 \cdots k_n} x_1^{k_1} \cdots x_n^{k_n}$$

se llama polinomio de n variables $x_1, \ldots, x_n$. Un polinomio es continuo en todo

$\mathbf{R}^n$ debido a que es una suma finita de productos de campos escalares continuos en todo $\mathbf{R}^n$. Por ejemplo, un polinomio de dos variables x e y, dado por

$$P(x, y) = \sum_{i=0}^{p} \sum_{j=0}^{q} c_{ij} x^i y^j$$

es continuo en todo punto (x, y) en $\mathbf{R}^2$.

EJEMPLO 5. *Continuidad de las funciones racionales.* Un campo escalar f dado por $f(\boldsymbol{x}) = P(\boldsymbol{x})/Q(\boldsymbol{x})$, donde P y Q son polinomios en los componentes de $\boldsymbol{x}$, se llama función racional. Tal función es continua en cada punto en donde $Q(\boldsymbol{x}) \neq 0$.

Pueden construirse otros ejemplos de funciones continuas con la ayuda del siguiente teorema, que está relacionado con la continuidad de las funciones compuestas.

TEOREMA 8.2. *Sean f y g dos funciones tales que la función compuesta $f \circ g$ está definida en $\boldsymbol{a}$, siendo*

$$(f \circ g)(\boldsymbol{x}) = f[g(\boldsymbol{x})].$$

Si g es continua en $\boldsymbol{a}$ y f es continua en $g(\boldsymbol{a})$, la función compuesta $f \circ g$ es continua en $\boldsymbol{a}$.

Demostración. Sean $\boldsymbol{y} = f(\boldsymbol{x})$ y $\boldsymbol{b} = g(\boldsymbol{a})$. Tenemos

$$f[g(\boldsymbol{x})] - f[g(\boldsymbol{a})] = f(\boldsymbol{y}) - f(\boldsymbol{b}).$$

Por hipótesis, $\boldsymbol{y} \to \boldsymbol{b}$ cuando $\boldsymbol{x} \to \boldsymbol{a}$, así que tenemos

$$\lim_{\|\boldsymbol{x}-\boldsymbol{a}\| \to 0} \|f[g(\boldsymbol{x})] - f[g(\boldsymbol{a})]\| = \lim_{\|\boldsymbol{y}-\boldsymbol{b}\| \to 0} \|f(\boldsymbol{y}) - f(\boldsymbol{b})\| = 0.$$

Por consiguiente $\lim_{\boldsymbol{x} \to \boldsymbol{a}} f[g(\boldsymbol{x})] = f[g(\boldsymbol{a})]$, así que $f \circ g$ es continua en $\boldsymbol{a}$.

EJEMPLO 6. El teorema anterior implica la continuidad de los campos escalares h, cuando $h(x, y)$ está dada por fórmulas tales como

$$\operatorname{sen}(x^2 y), \qquad \log(x^2 + y^2), \qquad \frac{e^{x+y}}{x+y}, \qquad \log[\cos(x^2 + y^2)].$$

En estos ejemplos las funciones son continuas en todos los puntos en los que están definidas. La primera lo es en todos los puntos del plano y la segunda en todos los puntos excepto en el origen. La tercera es continua en todos los puntos (x, y) en los que $x + y \neq 0$, y la cuarta en todos los puntos en los que $x^2 + y^2$ no sea un múltiplo impar de $\pi/2$. [El conjunto de los puntos (x, y) tales que $x^2 + y^2 = n\pi/2$, $n = 1, 3, 5, \ldots$, es una familia de circunferencias con centro en el origen.] Estos ejemplos demuestran que las discontinuidades de una función de dos variables pueden ser puntos aislados, curvas, o familias de curvas.

EJEMPLO 7. Una función de dos variables puede ser continua respecto a cada variable separadamente y en cambio ser discontinua como función de dos variables. Esto lo apreciamos en el ejemplo siguiente:

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{si} \quad (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

Para los puntos (x, y) del eje x tenemos $y = 0$ y $f(x, y) = f(x, 0) = 0$, así que la función tiene el valor constante 0 en todo el eje x. Por consiguiente, si ponemos $y = 0$ y consideramos f sólo como función de x, dicha f es continua en $x = 0$. Análogamente, f tiene el valor constante 0 en todos los puntos del eje y, así que si ponemos $x = 0$ y consideramos f como función sólo de y, f es continua en $y = 0$. No obstante, como función de las dos variables, f no es continua en el origen. En efecto, en cada punto de la recta $y = x$ (excepto en el origen) la función tiene el valor constante $\frac{1}{2}$ ya que $f(x, x) = x^2/(2x^2) = \frac{1}{2}$; puesto que existen puntos en esa recta tan próximos al origen como queramos y ya que $f(0, 0) \neq \frac{1}{2}$, la función no es continua en $(0, 0)$.

8.5 Ejercicios

Los ejercicios de esta sección se refieren a límites y a la continuidad de campos escalares definidos en subconjuntos del plano.

1. En cada uno de los siguientes ejemplos se define un campo escalar f mediante la ecuación dada para todos los puntos (x, y) del plano definidos por la expresión del segundo miembro. Determinar en cada ejemplo el conjunto de puntos (x, y) en los que f es continua.

a) $f(x, y) = x^4 + y^4 - 4x^2y^2$.

b) $f(x, y) = \log (x^2 + y^2)$.

c) $f(x, y) = \dfrac{1}{y} \cos x^2$.

d) $f(x, y) = \operatorname{tg} \dfrac{x^2}{y}$.

e) $f(x, y) = \operatorname{arc\,tg} \dfrac{y}{x}$.

f) $f(x, y) = \operatorname{arc\,sen} \dfrac{x}{\sqrt{x^2 + y^2}}$.

g) $f(x, y) = \arctan \dfrac{x + y}{1 - xy}$.

h) $f(x, y) = \dfrac{x}{\sqrt{x^2 + y^2}}$.

i) $f(x, y) = x^{(y^2)}$.

j) $f(x, y) = \arccos \sqrt{x/y}$.

2. Si $\lim_{(x,y)\to(a,b)} f(x, y) = L$, y si existen los dos límites uni-dimensionales

$$\lim_{x\to a} f(x, y) \qquad \text{y} \qquad \lim_{y\to b} f(x, y)$$

demostrar que

$$\lim_{x\to a}\, [\lim_{y\to b} f(x, y)] = \lim_{y\to b}\, [\lim_{x\to a} f(x, y)] = L.$$

Los dos límites de esta igualdad se llaman límites *iterados*; los ejercicios demuestran que la existencia del límite bidimensional y de los dos límites unidimensionales, implican la existencia e igualdad de los dos límites iterados. (El recíproco no siempre es cierto. En el ejercicio 4 se da un contraejemplo.)

3. Sea $f(x, y) = (x - y)/(x + y)$ si $x + y \neq 0$. Demostrar que

$$\lim_{x\to 0}\, [\lim_{y\to 0} f(x, y)] = 1 \qquad \text{pero que} \qquad \lim_{y\to 0}\, [\lim_{x\to 0} f(x, y)] = -1.$$

Utilizar ese resultado y el del ejercicio 2 para deducir que $f(x, y)$ no tiende a un límite cuando $(x, y) \to (0, 0)$.

4. Sea

$$f(x, y) = \frac{x^2y^2}{x^2y^2 + (x - y)^2} \qquad \text{siempre que} \quad x^2y^2 + (x - y)^2 \neq 0.$$

Demostrar que

$$\lim_{x\to 0}\, [\lim_{y\to 0} f(x, y)] = \lim_{y\to 0}\, [\lim_{x\to 0} f(x, y)] = 0$$

pero que $f(x, y)$ no tiende a un límite cuando $(x, y) \to (0, 0)$. [*Indicación*: Examinar f sobre la recta $y = x$.] Este ejemplo demuestra que el recíproco del ejercicio 2 no siempre es cierto.

5. Sea

$$f(x, y) = \begin{cases} x \operatorname{sen} \dfrac{1}{y} & \text{si} \quad y \neq 0, \\ 0 & \text{si} \quad y = 0. \end{cases}$$

Demostrar que $f(x, y) \to 0$ cuando $(x, y) \to (0, 0)$ pero que

$$\lim_{y\to 0}\, [\lim_{x\to 0} f(x, y)] \neq \lim_{x\to 0}\, [\lim_{y\to 0} f(x, y)].$$

Explicar por qué esto no contradice el ejercicio 2.

6. Si $(x, y) \neq (0, 0)$, pongamos $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$. Hallar el límite de $f(x,y)$ cuando $(x, y) \to (0, 0)$ a lo largo de la recta $y = mx$. ¿Es posible definir $f(0, 0)$ de modo que f sea continua en $(0, 0)$?
7. Sea $f(x, y) = 0$ si $y \leq 0$ o si $y \geq x^2$ y sea $f(x, y) = 1$ si $0 < y < x^2$. Demostrar que $f(x, y) \to 0$ cuando $(x, y) \to (0, 0)$ a lo largo de cualquier recta por el origen. Hallar una curva que pase por el origen a lo largo de la cual (salvo en el origen) $f(x, y)$ tiene el valor constante 1. ¿Es f continua en el origen?
8. Si $f(x, y) = [\operatorname{sen}(x^2 + y^2)]/(x^2 + y^2)$ cuando $(x, y) \neq (0, 0)$ ¿cómo debe definirse $f(0, 0)$ para que sea f continua en el origen?
9. Sea f un campo escalar continuo en un punto $\boldsymbol{a}$ interior a un conjunto S de $\mathbf{R}^n$. Si $f(\boldsymbol{a}) \neq 0$, demostrar que existe una n-bola $B(\boldsymbol{a})$ en la que f tiene el mismo signo que $f(\boldsymbol{a})$.

8.6 La derivada de un campo escalar respecto a un vector

Esta sección introduce las derivadas de campos escalares. En la sección 8.18 se discuten las derivadas de campos vectoriales.

Sea f un campo escalar definido en un conjunto S de $\mathbf{R}^n$, y sea $\boldsymbol{a}$ un punto interior a S. Deseamos estudiar la variación del campo cuando nos desplazamos desde $\boldsymbol{a}$ a un punto próximo. Por ejemplo, supongamos que $f(\boldsymbol{a})$ es la temperatura en un punto $\boldsymbol{a}$ en una habitación con calefacción y con una ventana abierta. Si nos movemos hacia la ventana la temperatura decrecerá, pero si nos acercamos al radiador de la calefacción aumentará. En general, la variación del campo dependerá de la dirección en la que nos movamos a partir de $\boldsymbol{a}$.

Supongamos que se representa esa dirección mediante otro vector $\boldsymbol{y}$. Esto es, supongamos que nos movemos desde $\boldsymbol{a}$ hacia otro punto $\boldsymbol{a} + \boldsymbol{y}$, siguiendo el segmento de recta que une $\boldsymbol{a}$ con $\boldsymbol{a} + \boldsymbol{y}$. Cada punto de este segmento tiene la forma $\boldsymbol{a} + h\boldsymbol{y}$, donde h es un número real. En la figura 8.3 se muestra un ejemplo. La distancia desde $\boldsymbol{a}$ hasta $\boldsymbol{a} + h\boldsymbol{y}$ es $\|h\boldsymbol{y}\| = |h|\ \|\boldsymbol{y}\|$.

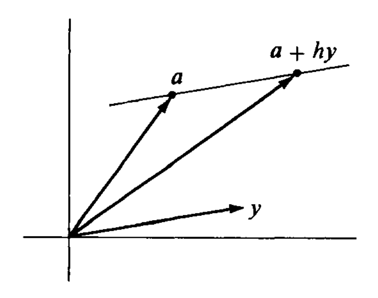


FIGURA 8.3 *El punto* $\boldsymbol{a} + h\boldsymbol{y}$ *está en la recta paralela a* $\boldsymbol{a}$ *que pasa por* $\boldsymbol{y}$.

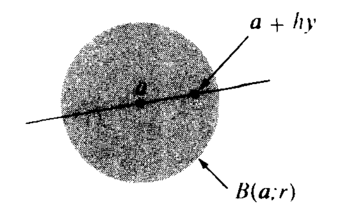


FIGURA 8.4 *El punto* $\boldsymbol{a} + h\boldsymbol{y}$ *está en la* n-*bola* $B(\boldsymbol{a}; r)$ *si* $\|h\boldsymbol{y}\| < r$.

Puesto que $\boldsymbol{a}$ es un punto interior de S, existe una n-bola $B(\boldsymbol{a}; r)$ situada enteramente en S. Si h se elige de manera que $|h| \, \|\boldsymbol{y}\| < r$, el segmento desde $\boldsymbol{a}$ hasta $\boldsymbol{a} + h\boldsymbol{y}$ estará en S. (Ver figura 8.4) Mantengamos $h \neq 0$ pero lo bastante pequeño para que $\boldsymbol{a} + h\boldsymbol{y} \in S$ y construyamos el cociente de diferencias

$$\frac{f(\boldsymbol{a} + h\boldsymbol{y}) - f(\boldsymbol{a})}{h}. \tag{8.3}$$

El numerador de este cociente pone de manifiesto el cambio de la función cuando nos desplazamos desde $\boldsymbol{a}$ a $\boldsymbol{a} + h\boldsymbol{y}$. El cociente se denomina a su vez el *promedio de variación* de f en el segmento de recta que une $\boldsymbol{a}$ con $\boldsymbol{a} + h\boldsymbol{y}$. Nos interesa el comportamiento de ese cociente cuando $h \to 0$. Esto nos lleva a la siguiente definición.

DEFINICIÓN DE LA DERIVADA DE UN CAMPO ESCALAR RESPECTO A UN VECTOR. *Dado un campo escalar $f\colon S \to \mathbf{R}$, donde $S \subseteq \mathbf{R}^n$. Sean $\boldsymbol{a}$ un punto interior a S e $\boldsymbol{y}$ un punto arbitrario de $\mathbf{R}^n$. La derivada de f en $\boldsymbol{a}$ con respecto a $\boldsymbol{y}$ se representa con el símbolo $f'(\boldsymbol{a}; \boldsymbol{y})$* y se define

$$f'(\boldsymbol{a}; \boldsymbol{y}) = \lim_{h \to 0} \frac{f(\boldsymbol{a} + h\boldsymbol{y}) - f(\boldsymbol{a})}{h} \tag{8.4}$$

cuando tal límite existe.

EJEMPLO 1. Si $\boldsymbol{y} = \boldsymbol{O}$, el cociente de diferencias (8.3) es 0 para todo $h \neq 0$, así que $f'(\boldsymbol{a}; \boldsymbol{O})$ existe siempre y es igual a 0.

EJEMPLO 2. *Derivada de una transformación lineal.* Si $f\colon S \to \mathbf{R}$ es lineal, entonces $f(\boldsymbol{a} + h\boldsymbol{y}) = f(\boldsymbol{a}) + hf(\boldsymbol{y})$ y el cociente de diferencias (8.3) es igual a $f(\boldsymbol{y})$ para todo $h \neq 0$. En este caso, $f'(\boldsymbol{a}; \boldsymbol{y})$ existe siempre y está dada por

$$f'(\boldsymbol{a}; \boldsymbol{y}) = f(\boldsymbol{y})$$

para todo $\boldsymbol{a}$ de S y todo $\boldsymbol{y}$ de $\mathbf{R}^n$. Dicho de otro modo, la derivada de una transformación lineal respecto a $\boldsymbol{y}$ es igual al valor de la función en $\boldsymbol{y}$.

Para estudiar el comportamiento de f sobre la recta que pasa por $\boldsymbol{a}$ y $\boldsymbol{a} + \boldsymbol{y}$ para $\boldsymbol{y} \neq \boldsymbol{O}$ introducimos la función

$$g(t) = f(\boldsymbol{a} + t\boldsymbol{y}).$$

El teorema que sigue relaciona las derivadas $g'(t)$ y $f'(\boldsymbol{a} + t\boldsymbol{y}; \boldsymbol{y})$.

TEOREMA 8.3. *Si $g(t) = f(\boldsymbol{a}+t\boldsymbol{y})$ y si una de las derivadas $g(t)$ o $f'(\boldsymbol{a}+t\boldsymbol{y};\boldsymbol{y})$ existe, entonces también existe la otra y coinciden,*

(8.5)
$$g'(t) = f'(\boldsymbol{a} + t\boldsymbol{y}; \boldsymbol{y}).$$

En particular, cuando $t = 0$ tenemos $g'(0) = f'(\boldsymbol{a}; \boldsymbol{y})$.

Demostración. Formando el cociente de diferencias para g, tenemos

$$\frac{g(t+h) - g(t)}{h} = \frac{f(\boldsymbol{a} + t\boldsymbol{y} + h\boldsymbol{y}) - f(\boldsymbol{a} + t\boldsymbol{y})}{h}$$

Haciendo que $h \to 0$ obtenemos (8.5).

EJEMPLO 3. Calcular $f'(\boldsymbol{a}; \boldsymbol{y})$ si $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ para todo $\boldsymbol{x}$ en $\mathbf{R}^n$.

Solución. Pongamos $g(t) = f(\boldsymbol{a} + t\boldsymbol{y}) = (\boldsymbol{a} + t\boldsymbol{y}) \cdot (\boldsymbol{a} + t\boldsymbol{y}) = \boldsymbol{a} \cdot \boldsymbol{a} + 2t\boldsymbol{a} \cdot \boldsymbol{y} + t^2\boldsymbol{y} \cdot \boldsymbol{y}$. Por consiguiente $g'(t) = 2\boldsymbol{a} \cdot \boldsymbol{y} + 2t\boldsymbol{y} \cdot \boldsymbol{y}$, así que $g'(0) = f'(\boldsymbol{a}; \boldsymbol{y}) = 2\boldsymbol{a} \cdot \boldsymbol{y}$.

Un corolario sencillo del Teorema 8.3 es el teorema del valor medio para campos escalares.

TEOREMA 8.4. TEOREMA DEL VALOR MEDIO PARA DERIVADAS DE CAMPOS ESCALARES. *Supongamos que existe la derivada $f'(\boldsymbol{a} + t\boldsymbol{y}; \boldsymbol{y})$ para cada t en el intervalo $0 \leqslant t \leqslant 1$. Entonces para un cierto número real θ en el intervalo abierto $0 < \theta < 1$ tenemos*

$$f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a}) = f'(\boldsymbol{z}; \boldsymbol{y}), \qquad \text{donde} \quad \boldsymbol{z} = \boldsymbol{a} + \theta\boldsymbol{y}.$$

Demostración. Pongamos $g(t) = f(\boldsymbol{a} + t\boldsymbol{y})$. Aplicando el teorema del valor medio uni-dimensional a g en el intervalo $[0, 1]$ tenemos

$$g(1) - g(0) = g'(\theta), \qquad \text{donde} \quad 0 < \theta < 1.$$

Puesto que $g(1) - g(0) = f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a})$ y $g'(\theta) = f'(\boldsymbol{a} + \theta\boldsymbol{y}; \boldsymbol{y})$, esto completa la demostración.

8.7 Derivadas direccionales y derivadas parciales

En el caso particular en el que $\boldsymbol{y}$ es un vector *unitario,* esto es, cuando $\|\boldsymbol{y}\| = 1$, la distancia entre $\boldsymbol{a}$ y $\boldsymbol{a} + h\boldsymbol{y}$ es $|h|$. En tal caso el cociente de diferen-

cias (8.3) representa el promedio de variación de f por *unidad de distancia* a lo largo del segmento de recta que une $\boldsymbol{a}$ con $\boldsymbol{a} + h\boldsymbol{y}$; la derivada $f'(\boldsymbol{a}; \boldsymbol{y})$ se denomina *derivada direccional.*

DEFINICIÓN DE DERIVADAS DIRECCIONAL Y PARCIAL. *Si* $\boldsymbol{y}$ *es un vector unitario, la derivada* $f'(\boldsymbol{a}; \boldsymbol{y})$ *se llama la derivada direccional de* f *en* $\boldsymbol{a}$ *en la dirección de* $\boldsymbol{y}$. *En particular, si* $\boldsymbol{y} = \boldsymbol{e}_k$ *(el k-ésimo vector coordenado unitario) la derivada direccional* $f'(\boldsymbol{a}; \boldsymbol{e}_k)$ *se denomina derivada parcial respecto a* $\boldsymbol{e}_k$ *y se representa también mediante el símbolo* $D_kf(\boldsymbol{a})$. *Así pues*

$$D_kf(\boldsymbol{a}) = f'(\boldsymbol{a}; \boldsymbol{e}_k).$$

Las siguientes notaciones se usan también para la derivada parcial $D_kf(\boldsymbol{a})$:

$$D_kf(a_1, \ldots, a_n), \qquad \frac{\partial f}{\partial x_k}(a_1, \ldots, a_n), \qquad \text{y} \qquad f'_{x_k}(a_1, \ldots, a_n).$$

Algunas veces la derivada f'_{x_k} se escribe sin el ápice f_{x_k}, o incluso más sencillamente f_k.

En $\mathbf{R}^2$ los vectores coordenados unidad se designan por $\boldsymbol{i}$ y $\boldsymbol{j}$. Si $\boldsymbol{a} = (a, b)$ las derivadas parciales $f'(\boldsymbol{a}; \boldsymbol{i})$ y $f'(\boldsymbol{a}; \boldsymbol{j})$ también se escriben

$$\frac{\partial f}{\partial x}(a, b) \qquad \text{y} \qquad \frac{\partial f}{\partial y}(a, b),$$

respectivamente. En $\mathbf{R}^3$, si $\boldsymbol{a} = (a, b, c)$ las derivadas parciales $D_1f(\boldsymbol{a}), D_2f(\boldsymbol{a})$, y $D_3f(\boldsymbol{a})$ se expresan poniendo

$$\frac{\partial f}{\partial x}(a, b, c), \qquad \frac{\partial f}{\partial y}(a, b, c), \qquad \text{y} \qquad \frac{\partial f}{\partial z}(a, b, c).$$

8.8 Derivadas parciales de orden superior

La derivación parcial origina nuevos campos escalares $D_1f, \ldots, D_nf$ a partir de un campo escalar dado f. Las derivadas parciales de $D_1f, \ldots, D_nf$ se llaman *derivadas parciales de segundo orden* de f. Para funciones de dos variables existen cuatro derivadas parciales de segundo orden, que se escriben así:

$$D_1(D_1f) = \frac{\partial^2 f}{\partial x^2}, \qquad D_1(D_2f) = \frac{\partial^2 f}{\partial x\, \partial y}, \qquad D_2(D_1f) = \frac{\partial^2 f}{\partial y\, \partial x}, \qquad D_2(D_2f) = \frac{\partial^2 f}{\partial y^2}.$$

Obsérvese que $D_1(D_2f)$ significa la derivada parcial de D_2f respecto a la primera variable. Algunas veces utilizamos la notación $D_{i,j}f$ para indicar la derivada parcial $D_i(D_jf)$. Por ejemplo, $D_{1,2}f = D_1(D_2f)$. También utilizamos la notación

$$\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

Ésta puede ser o no igual a la otra derivada parcial mixta,

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right).$$

En la sección 8.23 demostraremos que las dos derivadas parciales mixtas $D_1(D_2f)$ y $D_2(D_1f)$ son iguales en un punto si una de ellas es continua en un entorno del punto. También la sección 8.23 contiene un ejemplo en el que $D_1(D_2f) \neq D_2(D_1f)$ en un punto.

8.9 Ejercicios

1. Un campo escalar f está definido en $\mathbf{R}^n$ mediante la ecuación $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$, donde $\boldsymbol{a}$ es un vector constante. Calcular $f'(\boldsymbol{x}; \boldsymbol{y})$ cualesquiera que sean $\boldsymbol{x}$ e $\boldsymbol{y}$.
2. a) Resolver el ejercicio 1 cuando $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^4$.
   b) Tomar $n = 2$ en a) y hallar todos los puntos (x, y) para los cuales $f'(2\boldsymbol{i} + 3\boldsymbol{j}; x\boldsymbol{i} + y\boldsymbol{j}) = 6$.
   c) Tomar $n = 3$ en a) y hallar todos los puntos (x, y, z) para los cuales $f'(\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}; x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}) = 0$.
3. Sea $T: \mathbf{R}^n \to \mathbf{R}^n$ una transformación lineal dada. Calcular la derivada $f'(\boldsymbol{x}; \boldsymbol{y})$ para el campo escalar definido en $\mathbf{R}^n$ mediante la ecuación $f(\boldsymbol{x}) = \boldsymbol{x} \cdot T(\boldsymbol{x})$.

En cada uno de los ejercicios del 4 al 9, calcular todas las derivadas parciales de primer orden del campo escalar dado. En los ejercicios 8 y 9 los campos están definidos en $\mathbf{R}^n$.

4. $f(x, y) = x^2 + y^2 \operatorname{sen}(xy)$.
5. $f(x, y) = \sqrt{x^2 + y^2}$.
6. $f(x, y) = \dfrac{x}{\sqrt{x^2 + y^2}}$, $\quad (x, y) \neq (0, 0)$.
7. $f(x, y) = \dfrac{x + y}{x - y}$, $\quad x \neq y$.
8. $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$, $\quad \boldsymbol{a}$ fijo.
9. $f(\boldsymbol{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$, $\quad a_{ij} = a_{ji}$.

En cada uno de los ejercicios del 10 al 17, calcular todas las derivadas parciales de primer orden. En los ejercicios 10, 11 y 12 comprobar que las derivadas parciales mixtas $D_1(D_2f)$ y $D_2(D_1f)$ son iguales.

10. $f(x, y) = x^4 + y^4 - 4x^2y^2$.
11. $f(x, y) = \log(x^2 + y^2)$, $\quad (x, y) \neq (0, 0)$.
12. $f(x, y) = \dfrac{1}{y} \cos x^2$, $\quad y \neq 0$.
13. $f(x, y) = \tan(x^2/y)$, $\quad y \neq 0$.
14. $f(x, y) = \arctan(y/x)$, $\quad x \neq 0$.
15. $f(x, y) = \arctan \dfrac{x + y}{1 - xy}$, $\quad xy \neq 1$.
16. $f(x, y) = x^{(y^2)}$, $\quad x > 0$.
17. $f(x, y) = \arccos \sqrt{x/y}$, $\quad y \neq 0$.

18. Sea $v(r,t) = t^n e^{-r^2/(4t)}$. Hallar un valor de la constante n tal que v satisfaga la siguiente ecuación:

$$\frac{\partial v}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial v}{\partial r}\right)$$

19. Dada $z = u(x, y)e^{ax+by}$ y $\partial^2 u/\partial x\ \partial y) = 0$. Hallar valores de las constantes a y b tales que

$$\frac{\partial^2 z}{\partial x \partial y} - \frac{\partial z}{\partial x} - \frac{\partial z}{\partial y} + z = 0.$$

20. a) Suponer que $f'(\boldsymbol{x}; \boldsymbol{y}) = 0$ para todo $\boldsymbol{x}$ en una cierta n-bola $B(\boldsymbol{a})$ y para todo vector $\boldsymbol{y}$. Utilizar el teorema del valor medio para demostrar que f es constante en $B(\boldsymbol{a})$.
b) Suponer que $f'(\boldsymbol{x}; \boldsymbol{y}) = 0$ para un vector *fijo* $\boldsymbol{y}$ y todo $\boldsymbol{x}$ en $B(\boldsymbol{a})$. ¿Qué puede decirse de f en este caso?

21. Un conjunto S de $\mathbf{R}^n$ se llama *convexo* si para todo par de puntos $\boldsymbol{a}$ y $\boldsymbol{b}$ de S el segmento de recta que une $\boldsymbol{a}$ con $\boldsymbol{b}$ pertenece también a S; dicho de otro modo, $t\boldsymbol{a} + (1 - t)\boldsymbol{b} \in S$ para cada t del intervalo $0 \leq t \leq 1$.
a) Demostrar que toda n-bola es convexa.
b) Si $f'(\boldsymbol{x}; \boldsymbol{y}) = 0$ para todo $\boldsymbol{x}$ en un conjunto convexo abierto S y todo $\boldsymbol{y}$ de $\mathbf{R}^n$, demostrar que f es constante en S.

22. a) Demostrar que no existe un campo escalar f tal que $f'(\boldsymbol{a}; \boldsymbol{y}) > 0$ para un vector fijo $\boldsymbol{a}$ y todo vector no nulo $\boldsymbol{y}$.
b) Dar un ejemplo de un campo escalar f tal que $f'(\boldsymbol{x}; \boldsymbol{y}) > 0$ para un vector fijo $\boldsymbol{y}$ y todo vector $\boldsymbol{x}$.

8.10 Derivadas direccionales y continuidad

En la teoría uni-dimensional, la existencia de la derivada de una función f en un punto implica la continuidad en aquel punto. Esto se demuestra fácilmente eligiendo un $h \neq 0$ y escribiendo

$$f(a + h) - f(a) = \frac{f(a + h) - f(a)}{h} \cdot h.$$

Cuando $h \to 0$ el segundo miembro tiende al límite $f'(a) \cdot 0 = 0$ y por tanto $f(a + h) \to f(a)$. Esto prueba que la existencia de $f'(a)$ implica la continuidad de f en a.

Supongamos que aplicamos el mismo razonamiento a un campo escalar general. Supongamos que existe la derivada $f'(\boldsymbol{a}; \boldsymbol{y})$ para un cierto $\boldsymbol{y}$. Entonces si $h \neq 0$ podemos escribir

$$f(\boldsymbol{a} + h\boldsymbol{y}) - f(\boldsymbol{a}) = \frac{f(\boldsymbol{a} + h\boldsymbol{y}) - f(\boldsymbol{a})}{h} \cdot h.$$

Cuando $h \to 0$ el segundo miembro tiende al límite $f'(\boldsymbol{a}; \boldsymbol{y}) \cdot 0 = 0$; luego la existencia de $f'(\boldsymbol{a}; \boldsymbol{y})$ para un $\boldsymbol{y}$ dado implica que

$$\lim_{h \to 0} f(\boldsymbol{a} + h\boldsymbol{y}) = f(\boldsymbol{a})$$

para el mismo $\boldsymbol{y}$. Esto significa que $f(\boldsymbol{x}) \to f(\boldsymbol{a})$ cuando $\boldsymbol{x} \to \boldsymbol{a}$ *a lo largo de la recta de dirección* $\boldsymbol{y}$ *que pasa por* $\boldsymbol{a}$. Si $f'(\boldsymbol{a}; \boldsymbol{y})$ existe para todo *vector* $\boldsymbol{y}$, entonces $f(\boldsymbol{x}) \to f(\boldsymbol{a})$cuando $\boldsymbol{x} \to \boldsymbol{a}$ a lo largo de toda recta que pase por $\boldsymbol{a}$. Esto parece sugerir que f es continua en $\boldsymbol{a}$. Sin embargo, sorprende que esta conclusión no es necesariamente cierta. El ejemplo que sigue muestra un campo escalar que tiene derivada direccional en $\boldsymbol{O}$ en cualquier dirección pero que no es continuo en $\boldsymbol{O}$.

EJEMPLO. Sea f el campo escalar definido en $\mathbf{R}^2$ del modo siguiente:

$$f(x, y) = \frac{xy^2}{x^2 + y^4} \quad \text{si} \quad x \neq 0, \qquad f(0, y) = 0.$$

Sea $\boldsymbol{a} = (0, 0)$ e $\boldsymbol{y} = (a, b)$ cualquier vector. Si $a \neq 0$ y $h \neq 0$ tenemos

$$\frac{f(\boldsymbol{a} + h\boldsymbol{y}) - f(\boldsymbol{a})}{h} = \frac{f(h\boldsymbol{y}) - f(\boldsymbol{O})}{h} = \frac{f(ha, hb)}{h} = \frac{ab^2}{a^2 + h^2b^4}.$$

Haciendo que $h \to 0$ encontramos $f'(\boldsymbol{O}; \boldsymbol{y}) = b^2/a$. Si $\boldsymbol{y} = (0, b)$ encontramos, en forma parecida, que $f'(\boldsymbol{O}; \boldsymbol{y}) = 0$. Por consiguiente $f'(\boldsymbol{O}; \boldsymbol{y})$ existe para todas las direcciones $\boldsymbol{y}$. También, $f(\boldsymbol{x}) \to 0$ cuando $\boldsymbol{x} \to \boldsymbol{O}$ a lo largo de cualquier recta que pase por el origen. Sin embargo, en cada punto de la parábola $x = y^2$ (excepto en el origen) la función f tiene el valor $\frac{1}{2}$. Puesto que existen tales puntos tan próximos al origen como queramos y que $f(\boldsymbol{O}) = 0$, la función f no es continua en $\boldsymbol{O}$.

El ejemplo anterior prueba que la existencia de *todas* las derivadas direccionales en un punto no implican la continuidad en él. Por esta razón, las derivadas direccionales no constituyen una extensión satisfactoria del concepto uni-dimensional de derivada. Existe una generalización más conveniente que implica la continuidad y, al propio tiempo, nos permite extender los principales teoremas de la teoría de la derivación en una dimensión al caso de mayor número de dimensiones. Esa es la llamada *diferencial total* o simplemente *diferencial.*

8.11 La diferencial

Recordemos que en el caso uni-dimensional una función f que tiene derivada en a puede ser aproximada en un entorno de a mediante un polinomio de

Taylor de primer grado. Si existe $f'(a)$ designemos con $E(a, h)$ la diferencia

$$(8.6) \qquad E(a, h) = \frac{f(a+h) - f(a)}{h} - f'(a) \qquad \text{si} \quad h \neq 0.$$

y definamos $E(a, 0) = 0$. De (8.6) obtenemos la fórmula

$$f(a+h) = f(a) + f'(a)h + hE(a, h),$$

válida también para $h = 0$. Ésta es la fórmula de Taylor de primer orden para aproximar $f(a+h) - f(a)$ por medio de $f'(a)h$. El error cometido es $hE(a, h)$. De (8.6) resulta que $E(a, h) \to 0$ cuando $h \to 0$. Por consiguiente el error $hE(a, h)$ es de orden menor que h para valor pequeños de h.

Esta propiedad de aproximar una función diferenciable mediante una función lineal sugiere un método de extender el concepto de diferenciabilidad al caso de un número cualquiera de dimensiones.

Sea f: $S \to \mathbf{R}$ un campo escalar definido en un conjunto S de $\mathbf{R}^n$. Sean $\boldsymbol{a}$ un punto interior a S y $B(\boldsymbol{a}; r)$ una n-bola contenida en S. Sea $\boldsymbol{v}$ un vector tal que $\|\boldsymbol{v}\| < r$, de modo que $\boldsymbol{a} + \boldsymbol{v} \in B(\boldsymbol{a}; r)$.

DEFINICIÓN DE CAMPO ESCALAR DIFERENCIABLE. *Decimos que f es diferenciable en $\boldsymbol{a}$ si existe una transformación lineal*

$$T_a : \mathbf{R}^n \to \mathbf{R}$$

de $\mathbf{R}^n$ en $\mathbf{R}$, y una función escalar $E(\boldsymbol{a}, \boldsymbol{v})$ tal que

$$(8.7) \qquad f(\boldsymbol{a} + \boldsymbol{v}) = f(\boldsymbol{a}) + T_{\boldsymbol{a}}(\boldsymbol{v}) + \|\boldsymbol{v}\| \, E(\boldsymbol{a}, \boldsymbol{v}),$$

para $\|\boldsymbol{v}\| < r$ de manera que $E(\boldsymbol{a}, \boldsymbol{v}) \to 0$ cuando $\|\boldsymbol{v}\| \to 0$. La transformación lineal $T_{\boldsymbol{a}}$ se llama diferencial de f en $\boldsymbol{a}$.

Observación: La diferencial $T_{\boldsymbol{a}}$ es una transformación lineal, no un número. El valor $T_{\boldsymbol{a}}(\boldsymbol{v})$ es un número real; está definido para todo punto $\boldsymbol{v}$ de $\mathbf{R}^n$. La diferencial fue introducida por W. H. Young en 1908 y por M. Fréchet en 1911 en forma más general.

La ecuación (8.7), válida para $\|\boldsymbol{v}\| < r$, se llama *fórmula de Taylor de primer orden* para $f(\boldsymbol{a} + \boldsymbol{v})$. Nos proporciona una aproximación lineal, $T_{\boldsymbol{a}}(\boldsymbol{v})$, para la diferencia $f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a})$. El error en la aproximación es $\|\boldsymbol{v}\| \, E(\boldsymbol{a}, \boldsymbol{v})$, que es de orden más pequeño que $\|\boldsymbol{v}\|$ cuando $\|\boldsymbol{v}\| \to 0$; esto es, $E(\boldsymbol{a}, \boldsymbol{v}) = o(\|\boldsymbol{v}\|)$ cuando $\|\boldsymbol{v}\| \to 0$.

El teorema que sigue demuestra que si la diferencial existe, es única. Asimismo nos dice cómo calcular $T_a(y)$ para todo y de $\mathbf{R}^n$.

TEOREMA 8.5. *Si f es diferenciable en a con diferencial T_a, entonces existe la derivada $f'(a; y)$ para todo y de $\mathbf{R}^n$ y tenemos*

$$T_a(y) = f'(a; y). \tag{8.8}$$

Además, $f'(a; y)$ es una combinación lineal de los componentes de y. Efectivamente, si $y = (y_1, \ldots, y_n)$, tenemos

$$f'(a; y) = \sum_{k=1}^{n} D_k f(a) y_k. \tag{8.9}$$

Demostración. La ecuación (8.8) es trivial si $y = O$ puesto que $T_a(O) = 0$ y $f'(a; O) = 0$. Por consiguiente podemos suponer que $y \neq O$.

Puesto que f es diferenciable en a tenemos una fórmula de Taylor,

$$f(a + v) = f(a) + T_a(v) + \|v\| E(a, v) \tag{8.10}$$

para $\|v\| < r$ para algún $r > 0$, y donde $E(a, v) \to 0$ cuando $\|v\| \to 0$. En esta fórmula tomemos $v = hy$, siendo $h \neq 0$ y $|h| \ \|y\| < r$. Entonces $\|v\| < r$. Puesto que T_a es lineal es $T_a(v) = T_a(hy) = hT_a(y)$. Por lo tanto (8.10) nos da

$$\frac{f(a + hy) - f(a)}{h} = T_a(y) + \frac{|h| \ \|y\|}{h} E(a, v). \tag{8.11}$$

Ya que $\|v\| \to 0$ cuando $h \to 0$ y ya que $|h|/h = \pm 1$, el segundo miembro de (8.11) tiende al límite $T_a(y)$ cuando $h \to 0$. Por consiguiente el primer miembro tiene el mismo límite. Esto demuestra (8.8).

Para deducir (8.9) utilizamos la linealidad de T_a. Si $y = (y_1, \ldots, y_n)$ tenemos $y = \sum_{k=1}^{n} y_k e_k$, luego

$$T_a(y) = T_a\left(\sum_{k=1}^{n} y_k e_k\right) = \sum_{k=1}^{n} y_k T_a(e_k) = \sum_{k=1}^{n} y_k f'(a; e_k) = \sum_{k=1}^{n} y_k D_k f(a).$$

8.12 Gradiente de un campo escalar

La fórmula del teorema 8.5, que expresa $f'(a; y)$ como una combinación lineal de los componentes de y, puede escribirse como un producto escalar,

$$f'(a; y) = \sum_{k=1}^{n} D_k f(a) y_k = \nabla f(a) \cdot y,$$

donde $\nabla f(\boldsymbol{a})$ es el vector cuyos componentes son las derivadas parciales de f en $\boldsymbol{a}$,

$$\nabla f(\boldsymbol{a}) = (D_1 f(\boldsymbol{a}), \ldots, D_n f(\boldsymbol{a})).$$

Éste es el llamado *gradiente* de f. El gradiente ∇f es un campo vectorial definido en cada punto $\boldsymbol{a}$ en el que existen las derivadas parciales $D_1 f(\boldsymbol{a}), \ldots, D_n f(\boldsymbol{a})$. También empleamos la notación $\operatorname{grad} f$ en lugar de ∇f. El símbolo ∇ se lee «nabla».

La fórmula de Taylor de primer orden (8.10) puede ahora escribirse en la forma

$$f(\boldsymbol{a} + \boldsymbol{v}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} + \|\boldsymbol{v}\| \, E(\boldsymbol{a}, \boldsymbol{v}), \tag{8.12}$$

en donde $E(\boldsymbol{a}, \boldsymbol{v}) \to 0$ cuando $\|\boldsymbol{v}\| \to 0$. En esta forma se parece a la fórmula de Taylor unidimensional desempeñando el vector gradiente $\nabla f(\boldsymbol{a})$ el papel de la derivada $f'(\boldsymbol{a})$.

A partir de la fórmula de Taylor podemos deducir fácilmente que la diferenciabilidad implica la continuidad.

TEOREMA 8.6. *Si un campo escalar f es diferenciable en $\boldsymbol{a}$, entonces f es continua en $\boldsymbol{a}$.*

Demostración. De (8.12) resulta

$$|f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a})| = |\nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} + \|\boldsymbol{v}\| \, E(\boldsymbol{a}, \boldsymbol{v})|.$$

Aplicando la desigualdad triangular y la de Cauchy-Schwarz encontramos

$$0 \le |f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a})| \le \|\nabla f(\boldsymbol{a})\| \, \|\boldsymbol{v}\| + \|\boldsymbol{v}\| \, |E(\boldsymbol{a}, \boldsymbol{v})|.$$

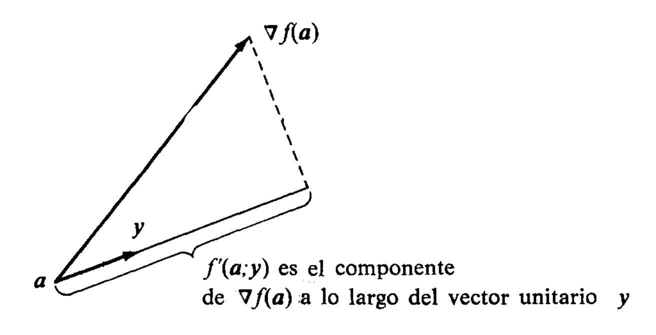


FIGURA 8.5 *Relación geométrica de la derivada direccional con el vector gradiente.*

Esto demuestra que $f(\boldsymbol{a}+\boldsymbol{v}) \to f(\boldsymbol{a})$ cuando $\|\boldsymbol{v}\| \to 0$, así que f es continua en $\boldsymbol{a}$.

Cuando $\boldsymbol{y}$ es un vector *unitario* la derivada direccional $f'(\boldsymbol{a};\boldsymbol{y})$ tiene una sencilla relación geométrica con el vector gradiente. Supongamos que $\nabla f(\boldsymbol{a}) \neq \boldsymbol{O}$ y designemos con θ el ángulo formado por $\nabla f(\boldsymbol{a})$ e $\boldsymbol{y}$. Tenemos entonces

$$f'(\boldsymbol{a};\boldsymbol{y}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{y} = \|\nabla f(\boldsymbol{a})\| \, \|\boldsymbol{y}\| \cos\theta = \|\nabla f(\boldsymbol{a})\| \cos\theta .$$

Esto nos dice que la derivada direccional es el componente del vector gradiente en la dirección del $\boldsymbol{y}$. La figura 8.5 muestra los vectores $\nabla f(\boldsymbol{a})$ e $\boldsymbol{y}$ aplicados al punto $\boldsymbol{a}$. La derivada alcanza el valor máximo cuando $\cos\theta = 1$, esto es, cuando $\boldsymbol{y}$ tiene la misma dirección que $\nabla f(\boldsymbol{a})$. Además, este máximo es igual a la longitud del vector gradiente. Cuando $\nabla f(\boldsymbol{a})$ es ortogonal a $\boldsymbol{y}$, la derivada direccional $f'(\boldsymbol{a};\boldsymbol{y})$ es 0. En el espacio de dos dimensiones se escribe con frecuencia

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x}\boldsymbol{i} + \frac{\partial f(x, y)}{\partial y}\boldsymbol{j} .$$

En tres dimensiones la fórmula correspondiente es

$$\nabla f(x, y, z) = \frac{\partial f(x, y, z)}{\partial x}\boldsymbol{i} + \frac{\partial f(x, y, z)}{\partial y}\boldsymbol{j} + \frac{\partial f(x, y, z)}{\partial z}\boldsymbol{k} .$$

8.13 Condición suficiente de diferenciabilidad

Si f es diferenciable en $\boldsymbol{a}$, existen todas las derivadas parciales $D_1 f(\boldsymbol{a}), \ldots, D_n f(\boldsymbol{a})$. No obstante, la existencia de todas esas derivadas no implica necesariamente que f sea diferenciable en $\boldsymbol{a}$. La siguiente función proporciona un contraejemplo

$$f(x, y) = \frac{xy^2}{x^2 + y^4} \quad \text{si} \quad x \neq 0, \qquad f(0, y) = 0,$$

que ya se discutió en la sección 8.10. Para esa función existen las dos derivadas parciales $D_1 f(\boldsymbol{O})$ y $D_2 f(\boldsymbol{O})$ pero f no es continua en $\boldsymbol{O}$, y por tanto no puede ser diferenciable en $\boldsymbol{O}$.

El teorema que sigue demuestra que la existencia de derivadas parciales *continuas* en un punto implica la diferenciabilidad en dicho punto.

TEOREMA 8.7. CONDICIÓN SUFICIENTE DE DIFERENCIABILIDAD. *Si existen las derivadas parciales $D_1 f, \ldots, D_n f$ en una cierta n-bola $B(\boldsymbol{a})$ y son continuas en $\boldsymbol{a}$, entonces f es diferenciable en $\boldsymbol{a}$.*

Observación: Un campo escalar que satisfaga las hipótesis 8.7 se le llama *diferenciable con continuidad* en $\boldsymbol{a}$.

Demostración. $T_{\boldsymbol{a}}(\boldsymbol{v})$ sólo puede ser $\nabla f(\boldsymbol{a}) \cdot \boldsymbol{v}$. Queremos demostrar que

$$f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} + \|\boldsymbol{v}\| \, E(\boldsymbol{a}, \boldsymbol{v}),$$

donde $E(\boldsymbol{a}, \boldsymbol{v}) \to 0$ cuando $\|\boldsymbol{v}\| \to 0$. Esto probará el teorema.

Pongamos $\lambda = \|\boldsymbol{v}\|$. Entonces $\boldsymbol{v} = \lambda \boldsymbol{u}$, donde $\|\boldsymbol{u}\| = 1$. Mantengamos λ lo bastante pequeño para que $\boldsymbol{a} + \boldsymbol{v}$ esté en la bola $B(\boldsymbol{a})$ en la que existen las derivadas parciales $D_1 f, \ldots, D_n f$. Expresando $\boldsymbol{u}$ en función de sus componentes tenemos

$$\boldsymbol{u} = u_1 \boldsymbol{e}_1 + \cdots + u_n \boldsymbol{e}_n,$$

donde $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ son los vectores coordenados unitarios. Escribamos ahora la diferencia $f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a})$ en forma de suma telescópica,

$$(8.13) \qquad f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a}) = f(\boldsymbol{a} + \lambda \boldsymbol{u}) - f(\boldsymbol{a}) = \sum_{k=1}^{n} \{f(\boldsymbol{a} + \lambda \boldsymbol{v}_k) - f(\boldsymbol{a} + \lambda \boldsymbol{v}_{k-1})\},$$

en la que $\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ son vectores cualesquiera de $\mathbf{R}^n$ tales que $\boldsymbol{v}_0 = \boldsymbol{O}$ y $\boldsymbol{v}_n = \boldsymbol{u}$. Elijamos esos vectores de modo que satisfagan la relación de recurrencia $\boldsymbol{v}_k = \boldsymbol{v}_{k-1} + u_k \boldsymbol{e}_k$. Esto es, tomemos

$$\boldsymbol{v}_0 = \boldsymbol{O}, \qquad \boldsymbol{v}_1 = u_1 \boldsymbol{e}_1, \qquad \boldsymbol{v}_2 = u_1 \boldsymbol{e}_1 + u_2 \boldsymbol{e}_2, \quad \ldots, \quad \boldsymbol{v}_n = u_1 \boldsymbol{e}_1 + \cdots + u_n \boldsymbol{e}_n.$$

Entonces el k-ésimo término de la suma (8.13) se transforma en

$$f(\boldsymbol{a} + \lambda \boldsymbol{v}_{k-1} + \lambda u_k \boldsymbol{e}_k) - f(\boldsymbol{a} + \lambda \boldsymbol{v}_{k-1}) = f(\boldsymbol{b}_k + \lambda u_k \boldsymbol{e}_k) - f(\boldsymbol{b}_k),$$

donde $\boldsymbol{b}_k = \boldsymbol{a} + \lambda \boldsymbol{v}_{k-1}$. Los dos puntos $\boldsymbol{b}_k$ y $\boldsymbol{b}_k + \lambda u_k \boldsymbol{e}_k$ tan sólo difieren en su k-ésimo componente. Por consiguiente podemos aplicar el teorema del valor medio del cálculo diferencial y escribir

$$(8.14) \qquad f(\boldsymbol{b}_k + \lambda u_k \boldsymbol{e}_k) - f(\boldsymbol{b}_k) = \lambda u_k \, D_k f(\boldsymbol{c}_k),$$

perteneciendo $\boldsymbol{c}_k$ al segmento de recta que une $\boldsymbol{b}_k$ a $\boldsymbol{b}_k + \lambda u_k \boldsymbol{e}_k$. Obsérvese que $\boldsymbol{b}_k \to \boldsymbol{a}$ y por tanto $\boldsymbol{c}_k \to \boldsymbol{a}$ cuando $\lambda \to 0$.

Aplicando (8.14) en (8.13) obtenemos

$$f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a}) = \lambda \sum_{k=1}^{n} D_k f(\boldsymbol{c}_k) u_k .$$

Pero $\nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} = \lambda \nabla f(\boldsymbol{a}) \cdot \boldsymbol{u} = \lambda \sum_{k=1}^{n} D_k f(\boldsymbol{a}) u_k$, así que

$$f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a}) - \nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} = \lambda \sum_{k=1}^{n} \{D_k f(\boldsymbol{c}_k) - D_k f(\boldsymbol{a})\} u_k = \|\boldsymbol{v}\| \, E(\boldsymbol{a}, \boldsymbol{v}),$$

donde

$$E(\boldsymbol{a}, \boldsymbol{v}) = \sum_{k=1}^{n} \{D_k f(\boldsymbol{c}_k) - D_k f(\boldsymbol{a})\} u_k .$$

Puesto que $\boldsymbol{c}_k \to \boldsymbol{a}$ cuando $\|\boldsymbol{v}\| \to 0$, y puesto que cada derivada parcial $D_k f$ es continua en $\boldsymbol{a}$, vemos que $E(\boldsymbol{a}, \boldsymbol{v}) \to 0$ cuando $\|\boldsymbol{v}\| \to 0$. Esto completa la demostración.

8.14 Ejercicios

1. Hallar el vector gradiente en cada punto en el que exista para los campos escalares definidos por las ecuaciones siguientes:
   a) $f(x, y) = x^2 + y^2 \operatorname{sen}(xy)$.
   b) $f(x, y) = e^x \cos y$.
   c) $f(x, y, z) = x^2 y^3 z^4$.
   d) $f(x, y, z) = x^2 - y^2 + 2z^2$.
   e) $f(x, y, z) = \log(x^2 + 2y^2 - 3z^2)$.
   f) $f(x, y, z) = x^{y^z}$.
2. Calcular las derivadas direccionales de los siguientes campos escalares en los puntos y direcciones que se indican:
   a) $f(x, y, z) = x^2 + 2y^2 + 3z^2$ en $(1, 1, 0)$ en la dirección de $\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}$.
   b) $f(x, y, z) = (x/y)^z$ en $(1, 1, 1)$ en la dirección de $2\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}$.
3. Hallar los puntos (x, y) y las direcciones para las que la derivada direccional de $f(x, y) = 3x^2 + y^2$ tiene el valor máximo, si (x, y) está en el círculo $x^2 + y^2 = 1$.
4. Un campo escalar diferenciable f tiene, en el punto $(1, 2)$ las derivadas direccionales $+2$ en dirección al punto $(2, 2)$ y -2 en dirección al punto $(1, 1)$. Determinar el vector gradiente en $(1, 2)$ y calcular la derivada direccional en dirección al punto $(4,6)$.
5. Hallar los valores de las constantes a, b y c tales que la derivada direccional de $f(x, y, z) = axy^2 + byz + cz^2x^3$ en el punto $(1, 2, -1)$ tenga el valor máximo 64 en la dirección paralela al eje z.
6. Dado un campo escalar diferenciable en un punto $\boldsymbol{a}$ de $\mathbf{R}^2$. Supongamos que $f'(\boldsymbol{a}; \boldsymbol{y}) = 1$ y $f'(\boldsymbol{a}; \boldsymbol{z}) = 2$, siendo $\boldsymbol{y} = 2\boldsymbol{i} + 3\boldsymbol{j}$ y $\boldsymbol{z} = \boldsymbol{i} + \boldsymbol{j}$. Hacer un gráfico mostrando el conjunto de todos los puntos (x, y) para los que $f'(\boldsymbol{a}; x\boldsymbol{i} + y\boldsymbol{j}) = 6$. Calcular también el gradiente $\nabla f(\boldsymbol{a})$.
7. Sean f y g dos campos escalares diferenciables en un conjunto abierto S. Deducir las siguientes propiedades del gradiente:
   a) $\operatorname{grad} f = \boldsymbol{O}$ si f es constante en S.
   b) $\operatorname{grad}(f + g) = \operatorname{grad} f + \operatorname{grad} g$.
   c) $\operatorname{grad}(cf) = c \operatorname{grad} f$ si c es constante.
   d) $\operatorname{grad}(fg) = f \operatorname{grad} g + g \operatorname{grad} f$.
   e) $\operatorname{grad}\left(\dfrac{f}{g}\right) = \dfrac{g \operatorname{grad} f - f \operatorname{grad} g}{g^2}$ en los puntos en los que $g \neq 0$.
8. En $\mathbf{R}^3$ consideremos $\boldsymbol{r}(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$, y sea $r(x, y, z) = \|\boldsymbol{r}(x, y, z)\|$.

a) Demostrar que $\nabla r(x, y, z)$ es un vector unidad en la dirección de $\boldsymbol{r}$ (x, y, z).
b) Demostrar que $\nabla(r^n) = nr^{n-2}\boldsymbol{r}$ si n es un entero positivo.
[*Indicación:* Utilizar el ejercicio 7 d).]
c) ¿Es válida la fórmula del apartado b) cuando n es entero negativo o cero?
d) Hallar un campo escalar f tal que $\nabla f = \boldsymbol{r}$.

9. Supongamos que f es diferenciable en cada punto de una n-bola $B(\boldsymbol{a})$. Si $f'(\boldsymbol{x}; \boldsymbol{y}) = 0$ para n vectores independientes $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ y para todo $\boldsymbol{x}$ en $B(\boldsymbol{a})$, demostrar que f es constante en $B(\boldsymbol{a})$.

10. Supongamos que f es diferenciable en cada punto de una n-bola $B(\boldsymbol{a})$.
a) Si $\nabla f(\boldsymbol{x}) = \boldsymbol{O}$ para todo $\boldsymbol{x}$ de $B(\boldsymbol{a})$, demostrar que f es constante en $B(\boldsymbol{a})$.
b) Si $f(\boldsymbol{x}) \leq f(\boldsymbol{a})$ para todo $\boldsymbol{x}$ de $B(\boldsymbol{a})$, demostrar que $\nabla f(\boldsymbol{a}) = \boldsymbol{O}$.

11. Considerar las seis proposiciones que siguen relativas a un campo escalar $f: S \to \mathbf{R}$, siendo $S \subseteq \mathbf{R}^n$ y $\boldsymbol{a}$ un punto interior a S.
a) f es continuo en $\boldsymbol{a}$.
b) f es diferenciable en $\boldsymbol{a}$.
c) $f'(\boldsymbol{a}; \boldsymbol{y})$ existe para todo $\boldsymbol{y}$ de $\mathbf{R}^n$.
d) Existen todas las derivadas parciales de f en un entorno de $\boldsymbol{a}$ y son continuas en $\boldsymbol{a}$.
e) $\nabla f(\boldsymbol{a}) = \boldsymbol{O}$.
f) $f(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}\|$ para todo $\boldsymbol{x}$ de $\mathbf{R}^n$.

En una tabla parecida a la indicada aquí, marcar con una T el cuadrado correspondiente si la proposición de la fila (x) implica siempre la proposición de la columna (y). Por ejemplo, si (a) implica siempre (b), poner T en el segundo cuadrado de la primera fila. La diagonal principal ha sido ya marcada.

	a	b	c	d	e	f
a	T					
b		T				
c			T			
d				T		
e					T	
f						T

8.15 Regla de la cadena para derivadas de campos escalares

En la teoría de la derivación en una dimensión, la regla de la cadena nos permite calcular la derivada de una función compuesta $g(t) = f[r(t)]$ mediante la fórmula

$$g'(t) = f'[r(t)] \cdot r'(t).$$

Esta sección nos proporciona una extensión de la fórmula cuando f se reemplaza por un campo escalar definido en un conjunto del espacio de dimensión n, y $\boldsymbol{r}$ por una función vectorial de una variable real cuyos valores están en el dominio de f.

Más adelante extenderemos aún más la fórmula para incluir el caso en el que f y r son ambas campos vectoriales.

Resulta fácil imaginarse casos en los que puede presentarse la composición de un campo escalar y un campo vectorial. Por ejemplo, supongamos que $f(x)$ mide la temperatura en un punto x de un sólido tri-dimensional, y supongamos que queremos conocer cómo cambia la temperatura cuando el punto x se mueve a lo largo de una curva C situada en el sólido. Si la curva es descrita por una función vectorial r definida en un intervalo $[a, b]$, podemos introducir una nueva función g mediante la fórmula

$$g(t) = f[r(t)] \qquad si \quad a \leq t \leq b.$$

Esta función compuesta g expresa la temperatura como función del parámetro t, y su derivada $g'(t)$ mide la variación de la temperatura a lo largo de la curva. La siguiente extensión de la regla de la cadena nos permite calcular la derivada $g'(t)$ sin determinar $g(t)$ explícitamente.

TEOREMA 8.8. REGLA DE LA CADENA. *Sea f un campo escalar definido en en un conjunto abierto S de $\mathbf{R}^n$, y sea r una función vectorial que aplica un intervalo J de $\mathbf{R}^1$ en S. Definamos la función compuesta $g = f \circ r$ en J mediante la ecuación*

$$g(t) = f[r(t)] \qquad si \quad t \in J.$$

Sea t un punto de J en el que $r'(t)$ existe y supongamos que f es diferenciable en $r(t)$. Existe entonces $g'(t)$ y es igual al producto escalar

$$g'(t) = \nabla f(a) \cdot r'(t), \qquad donde \quad a = r(t). \tag{8.15}$$

Demostración. Pongamos $a = r(t)$, siendo t un punto de J en el que $r'(t)$ exista. Puesto que S es abierto existe una n-bola $B(a)$ situada en S. Tomemos $h \neq 0$ lo bastante pequeño para que $r(t + h)$ esté situada en $B(a)$, y pongamos $y = r(t + h) - r(t)$. Obsérvese que $y \to O$ cuando $h \to 0$. Tenemos ahora

$$g(t + h) - g(t) = f[r(t + h)] - f[r(t)] = f(a + y) - f(a).$$

Aplicando la fórmula de Taylor de primer orden a f tenemos

$$f(a + y) - f(a) = \nabla f(a) \cdot y + \|y\| \, E(a, y),$$

donde $E(a, y) \to 0$ cuando $\|y\| \to 0$. Ya que $y = r(t + h) - r(t)$ esto nos da

$$\frac{g(t + h) - g(t)}{h} = \nabla f(a) \cdot \frac{r(t + h) - r(t)}{h} + \frac{\|r(t + h) - r(t)\|}{h} E(a, y).$$

Haciendo que $h \to 0$ obtenemos (8.15).

EJEMPLO 1. *Derivada direccional a lo largo de una curva.* Cuando la función $\boldsymbol{r}$ describe una curva C, la derivada $\boldsymbol{r}'$ es el vector velocidad (tangente a la curva) y la derivada g' de la ecuación (8.15) es la derivada de f respecto al vector velocidad, suponiendo que $\boldsymbol{r}' \neq \boldsymbol{O}$. Si $\boldsymbol{T}(t)$ es un vector unitario en la dirección de $\boldsymbol{r}'(t)$ ($\boldsymbol{T}$ es el vector tangente unitario), el producto escalar $\nabla f[\boldsymbol{r}(t)] \cdot \boldsymbol{T}(t)$ se llama *derivada direccional de f a lo largo de la curva C* o *en la dirección de C.* Para una curva plana podemos escribir

$$\boldsymbol{T}(t) = \cos \alpha(t)\, \boldsymbol{i} + \cos \beta(t)\, \boldsymbol{j},$$

siendo $\alpha(t)$ y $\beta(t)$ los ángulos formados por el vector $\boldsymbol{T}(t)$ y los ejes x e y positivos; la derivada direccional de f a lo largo de C es en este caso

$$\nabla f[\boldsymbol{r}(t)] \cdot \boldsymbol{T}(t) = D_1 f[\boldsymbol{r}(t)] \cos \alpha(t) + D_2 f[\boldsymbol{r}(t)] \cos \beta(t).$$

Con frecuencia esta fórmula se escribe más simplemente así:

$$\nabla f \cdot \boldsymbol{T} = \frac{\partial f}{\partial x} \cos \alpha + \frac{\partial f}{\partial y} \cos \beta.$$

Algunos autores expresan la derivada direccional $\nabla f \cdot \boldsymbol{T}$ con el símbolo df/ds. Puesto que la derivada direccional a lo largo de C está definida en función de $\boldsymbol{T}$, su valor depende de la representación paramétrica elegida para C. Un cambio de la representación podría invertir la dirección de $\boldsymbol{T}$; lo que, a su vez, invertirá el signo de la derivada direccional.

EJEMPLO 2. Hallar la derivada direccional del campo escalar $f(x, y) = x^2 - 3xy$ a lo largo de la parábola $y = x^2 - x + 2$ en el punto $(1, 2)$.

Solución. En un punto cualquiera (x, y) el vector gradiente es

$$\nabla f(x, y) = \frac{\partial f}{\partial x} \boldsymbol{i} + \frac{\partial f}{\partial y} \boldsymbol{j} = (2x - 3y)\boldsymbol{i} - 3x\boldsymbol{j}.$$

En el punto $(1, 2)$ tenemos $\nabla f(1, 2) = -4\boldsymbol{i} - 3\boldsymbol{j}$. La parábola puede representarse paramétricamente mediante la ecuación vectorial $\boldsymbol{r}(t) = t\boldsymbol{i} + (t^2 - t + 2)\boldsymbol{j}$. Por lo tanto, $\boldsymbol{r}(1) = \boldsymbol{i} + 2\boldsymbol{j}$, $\boldsymbol{r}'(t) = \boldsymbol{i} + (2t - 1)\boldsymbol{j}$, y $\boldsymbol{r}'(1) = \boldsymbol{i} + \boldsymbol{j}$. Para esta representación de C el vector unitario tangente $\boldsymbol{T}(1)$ es $(\boldsymbol{i} + \boldsymbol{j})/\sqrt{2}$ y la derivada di-

reccional pedida es $\nabla f(1,2) \cdot \boldsymbol{T}(1) = -7/\sqrt{2}$.

EJEMPLO 3. Sean f un campo escalar no constante, diferenciable en todo el plano, y c una constante. Supongamos que la ecuación cartesiana $f(x, y) = c$ describa una curva C que tenga tangente en cada uno de sus puntos. Demostrar que f tiene las siguientes propiedades en cada punto de C:

a) El vector gradiente ∇f es normal a C.

b) La derivada direccional de f es cero a lo largo de C.

c) La derivada direccional de f tiene su valor máximo en la dirección normal a C.

Solución. Si $\boldsymbol{T}$ es un vector unitario tangente a C, la derivada direccional de f a lo largo de C es el producto escalar $\nabla f \cdot \boldsymbol{T}$. Este producto es cero si ∇f es perpendicular a $\boldsymbol{T}$, y alcanza su máximo valor si ∇f es paralela a $\boldsymbol{T}$. Por consiguiente las dos proposiciones b) y c) son consecuencias de a). Para demostrar a), consideremos una curva plana cualquiera Γ con una ecuación vectorial de la forma $\boldsymbol{r}(t) = X(t)\boldsymbol{i} + Y(t)\boldsymbol{j}$ e introduzcamos la función $g(t) = f[\boldsymbol{r}(t)]$. En virtud de la regla de la cadena tenemos $g'(t) = \nabla f[\boldsymbol{r}(t)] \cdot \boldsymbol{r}'(t)$. Cuando $\Gamma = C$, la función g tiene el valor constante c así que $g'(t) = 0$ si $\boldsymbol{r}(t) \in C$. Puesto que $g' = \nabla f \cdot \boldsymbol{r}'$, resulta que ∇f es perpendicular a $\boldsymbol{r}'$ en C; luego ∇f es normal a C.

8.16 Aplicaciones geométricas. Conjuntos de nivel. Planos tangentes

La regla de la cadena puede utilizarse para deducir propiedades geométricas

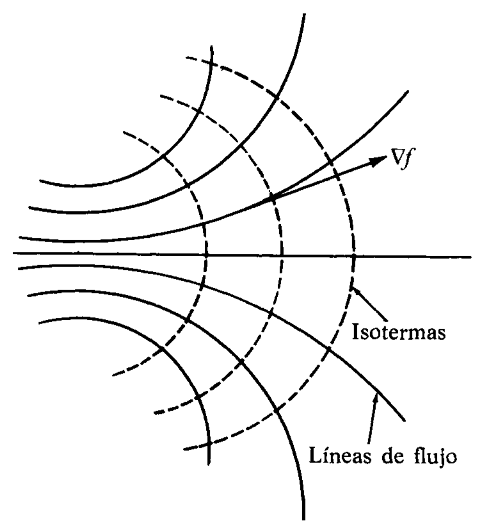


FIGURA 8.6 *Las curvas de trazos son isotermas: $f(x, y) = c$. El vector gradiente ∇f indica la dirección de las líneas de fuerza.*

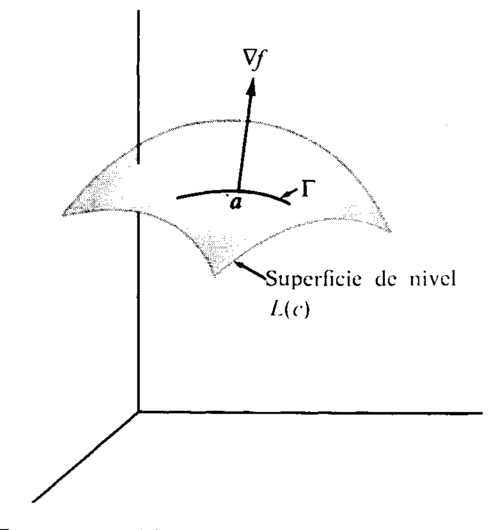


FIGURA 8.7 *El vector gradiente ∇f es normal a cada curva Γ situada en la superficie de nivel $f(x, y, z) = c$.*

del vector gradiente. Sea f un campo escalar definido en un conjunto S de $\mathbf{R}^n$ y consideremos aquellos puntos x de S para los que $f(x)$ tiene un valor constante, por ejemplo $f(x) = c$. Designemos ese conjunto por $L(c)$, de modo que

$$L(c) = \{x \mid x \in S \quad \text{y} \quad f(x) = c\}.$$

El conjunto $L(c)$ se llama *conjunto de nivel* de f. En $\mathbf{R}^2$, $L(c)$ se llama *curva de nivel*; en $\mathbf{R}^3$ *superficie de nivel*.

En muchas aplicaciones físicas se presentan familias de curvas de nivel. Por ejemplo, si $f(x, y)$ representa la temperatura en (x, y), las curvas de nivel de f (curvas de temperatura constante) se llaman *isotermas*. El flujo de calor tiene lugar en la dirección del cambio más rápido de la temperatura. Luego, en una hoja plana delgada el flujo de calor tiene lugar a lo largo de una familia de curvas ortogonales a las isotermas. Esas se llaman *líneas de flujo;* son las trayectorias ortogonales de las isotermas. Véase la figura 8.6.

Consideremos ahora un campo escalar f diferenciable en un conjunto abierto S de $\mathbf{R}^3$, y examinemos una de sus superficies de nivel, $L(c)$. Sea a un punto en esa superficie, y consideremos una curva Γ situada en la superficie y que pase por a, como está indicado en la figura 8.7. Demostraremos que el vector gradiente

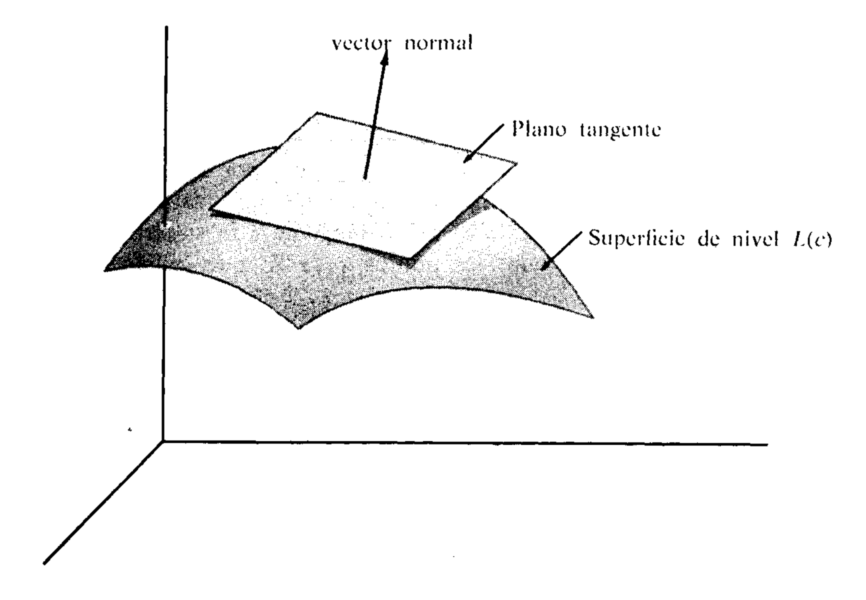


FIGURA 8.8 *El vector gradiente ∇f es normal al plano tangente a una superficie de nivel $f(x, y, z) = c$.*

$\nabla f(\boldsymbol{a})$ es normal a esa curva en $\boldsymbol{a}$. Esto es, demostraremos que $\nabla f(\boldsymbol{a})$ es perpendicular al vector tangente de Γ en $\boldsymbol{a}$. A tal fin supongamos que Γ está definida paramétricamente por una función vectorial derivable $\boldsymbol{r}$ definida en un cierto intervalo J de $\mathbf{R}^1$. Puesto que Γ está situada en la superficie de nivel $L(c)$, la función $\boldsymbol{r}$ satisface la ecuación

$$f[\boldsymbol{r}(t)] = c \text{ para todo } t \text{ de } J.$$

Si $g(t) = f[\boldsymbol{r}(t)]$ para t en J, la regla de la cadena establece que

$$g'(t) = \nabla f[\boldsymbol{r}(t)] \cdot \boldsymbol{r}'(t).$$

Puesto que g es constante en J, tenemos $g'(t) = 0$ en J. En particular, eligiendo t_1 de modo que $g(t_1) = \boldsymbol{a}$ encontramos que

$$\nabla f(\boldsymbol{a}) \cdot \boldsymbol{r}'(t_1) = 0.$$

Es decir, el gradiente de f en $\boldsymbol{a}$ es perpendicular al vector tangente $\boldsymbol{r}'(t_1)$ como se afirmó.

Tomemos ahora una familia de curvas en la superficie de nivel $L(c)$ que pasen por el punto $\boldsymbol{a}$. Según lo dicho en la discusión anterior, los vectores tangentes a todas esas curvas son perpendiculares al vector gradiente $\nabla f(\boldsymbol{a})$. Si éste no es cero, dichos vectores tangentes determinan un plano, y el gradiente $\nabla f(\boldsymbol{a})$ es normal a este plano. (Véase figura 8.8.) Tal plano se llama *plano tangente* a la superficie $L(c)$ en $\boldsymbol{a}$.

En el Volumen I se vio que un plano que pase por $\boldsymbol{a}$ con vector normal $\boldsymbol{N}$ está constituido por todos los puntos $\boldsymbol{x}$ de $\mathbf{R}^3$ que satisfacen la ecuación $\boldsymbol{N} \cdot (\boldsymbol{x} - \boldsymbol{a}) = 0$. Por consiguiente el plano tangente a la superficie de nivel $L(c)$ en $\boldsymbol{a}$ estará constituido por todos los puntos $\boldsymbol{x}$ de $\mathbf{R}^3$ que satisfagan

$$\nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}) = 0.$$

Para obtener una ecuación cartesiana de ese plano expresaremos $\boldsymbol{x}$, $\boldsymbol{a}$, y $\nabla f(\boldsymbol{a})$ en función de sus componentes. Escribiendo $\boldsymbol{x} = (x, y, z)$, $\boldsymbol{a} = (x_1, y_1, z_1)$, y

$$\nabla f(\boldsymbol{a}) = D_1 f(\boldsymbol{a})\boldsymbol{i} + D_2 f(\boldsymbol{a})\boldsymbol{j} + D_3 f(\boldsymbol{a})\boldsymbol{k},$$

obtenemos la ecuación cartesiana

$$D_1 f(\boldsymbol{a})(x - x_1) + D_2 f(\boldsymbol{a})(y - y_1) + D_3 f(\boldsymbol{a})(z - z_1) = 0.$$

A los campos escalares definidos en $\mathbf{R}^2$ se aplica una discusión parecida. En el ejemplo 3 de la sección anterior hemos demostrado que el vector gradiente $\nabla f(\boldsymbol{a})$ en un punto $\boldsymbol{a}$ de una curva de nivel es perpendicular al vector tangente a la curva en $\boldsymbol{a}$. Por consiguiente la recta tangente a la curva de nivel $L(c)$ en un punto $\boldsymbol{a} = (x_1, y_1)$ tiene la ecuación cartesiana

$$D_1 f(\boldsymbol{a})(x - x_1) + D_2 f(\boldsymbol{a})(y - y_1) = 0 .$$

8.17 Ejercicios

1. En este ejercicio podemos suponer la existencia y continuidad de todas las derivadas que se consideren. Las ecuaciones $u = f(x,y)$, $x = X(t)$, $y = Y(t)$ definen u como función de t, pongamos $u = F(t)$.
   a) Aplicar la regla de la cadena para demostrar que

   $$F'(t) = \frac{\partial f}{\partial x} X'(t) + \frac{\partial f}{\partial y} Y'(t),$$

   donde $\partial f/\partial x$ y $\partial f/\partial y$ se han calculado en $[X(t), Y(t)]$.
   b) En forma parecida, expresar la derivada segunda $F''(t)$ en función de las derivadas de f, X e Y. Recuérdese que las derivadas parciales de la fórmula del apartado a) son funciones compuestas dadas por

   $$\frac{\partial f}{\partial x} = D_1 f[X(t), Y(t)], \qquad \frac{\partial f}{\partial y} = D_2 f[X(t), Y(t)].$$

2. Teniendo en cuenta el ejercicio 1 calcular $F'(t)$ y $F''(t)$ en función de t en cada uno de los siguientes casos particulares:
   a) $f(x, y) = x^2 + y^2$, $X(t) = t$, $Y(t) = t^2$.
   b) $f(x, y) = e^{xy} \cos (xy^2)$, $X(t) = \cos t$, $Y(t) = \operatorname{sen} t$.
   c) $f(x, y) = \log [(1 + e^{x^2})/(1 + e^{y^2})]$, $X(t) = e$, $Y(t)^t = e^{-t}$.
3. Calcular en cada caso la derivada direccional de f en los puntos y direcciones que se indican:
   a) $f(x, y, z) = 3x - 5y + 2z$ en $(2, 2, 1)$ en la dirección de la normal exterior a la esfera $x^2 + y^2 + z^2 = 9$.
   b) $f(x, y, z) = x^2 - y^2$ en un punto cualquiera de la superficie $x^2 + y^2 + z^2 = 4$ en la dirección de la normal exterior en dicho punto.
   c) $f(x, y, z) = x^2 + y^2 - z^2$ en $(3, 4, 5)$ a lo largo de la curva de intersección de las dos superficies $2x^2 + 2y^2 - z^2 = 25$ y $x^2 + y^2 = z^2$.
4. a) Hallar un vector $V(x, y, z)$ normal a la superficie

   $$z = \sqrt{x^2 + y^2} + (x^2 + y^2)^{3/2}$$

en un punto cualquiera (x, y, z) de la superficie $(x, y, z) \neq (0, 0, 0)$.
b) Hallar el coseno del ángulo θ formado por el vector $V(x, y, z)$ y el eje z y determinar el límite de $\cos\theta$ cuando $(x, y, z) \to (0, 0, 0)$.

5. Las dos ecuaciones $e^u \cos v = x$ y $e^u \operatorname{sen} v = y$ definen u y v como funciones de x e y, sean éstas $u = U(x, y)$ y $v = V(x, y)$. Hallar fórmulas explícitas para $U(x, y)$ y $V(x, y)$, válidas para $x > 0$, y demostrar que los vectores gradientes $\nabla U(x, y)$ y $\nabla V(x, y)$ son perpendiculares en cada punto (x, y).
6. Sea $f(x, y) = \sqrt{|xy|}$.
a) Comprobar que $\partial f/\partial x$ y $\partial f/\partial y$ son cero ambas en el origen.
b) ¿Tiene la superficie $z = f(x, y)$ plano tangente en el origen? [*Indicación*: Considérese la sección producida en la superficie por el plano $x = y$.]
7. Si (x_0, y_0, z_0) es un punto de la superficie $z = xy$, las dos rectas $z = y_0 x$, $y = y_0$ y $z = x_0 y$, $x = x_0$ se cortan en (x_0, y_0, z_0) y están situadas en la superficie. Comprobar que el plano tangente a esta superficie en el punto (x_0, y_0, z_0) contiene a esas dos rectas.
8. Hallar la ecuación cartesiana del plano tangente a la superficie $xyz = a^3$ en un punto genérico (x_0, y_0, z_0). Demostrar que el volumen del tetraedro limitado por ese plano y los tres planos coordenados es $9a^3/2$.
9. Hallar un par de ecuaciones cartesianas para la recta que es tangente a las dos superficies $x^2 + y^2 + 2z^2 = 4$ y $z = e^{x-y}$ en el punto $(1, 1, 1)$.
10. Hallar una constante c tal que en todo punto de la intersección de las dos esferas

$$(x - c)^2 + y^2 + z^2 = 3 \qquad \text{y} \qquad x^2 + (y - 1)^2 + z^2 = 1$$

los planos tangentes correspondientes sean perpendiculares el uno al otro.

11. Si r_1 y r_2 son las distancias desde un punto (x, y) de una elipse a sus focos, demostrar que la ecuación $r_1 + r_2 =$ constante (que satisfacen esas distancias) implica la relación

$$\boldsymbol{T} \cdot \nabla(r_1 + r_2) = 0,$$

siendo $\boldsymbol{T}$ el vector unitario tangente a la curva. Interpretar geométricamente ese resultado, y con ello demostrar que la tangente forma ángulos iguales con las rectas que unen (x, y) a los focos.

12. Si $\nabla f(x, y, z)$ es siempre paralela a $x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$, demostrar que f debe tomar valores iguales en los puntos $(0, 0, a)$ y $(0, 0, -a)$.

8.18 Diferenciales de campos vectoriales

La teoría de la diferenciación para campos vectoriales es una extensión directa de la teoría análoga para campos escalares. Sea $f: S \to \mathbf{R}^m$ un campo vectorial definido en un subconjunto S de $\mathbf{R}^n$. Si $\boldsymbol{a}$ es un punto interior de S e $\boldsymbol{y}$ un vector cualquiera de $\mathbf{R}^n$ definimos la derivada $f'(\boldsymbol{a}; \boldsymbol{y})$ mediante la fórmula

$$f'(\boldsymbol{a}; \boldsymbol{y}) = \lim_{h \to 0} \frac{f(\boldsymbol{a} + h\boldsymbol{y}) - f(\boldsymbol{a})}{h},$$

siempre que tal límite exista. La derivada $f'(\boldsymbol{a}; \boldsymbol{y})$ es un vector de $\mathbf{R}^m$.

Designemos con f_k el k-ésimo componente de f. Observemos que la derivada $f'(\boldsymbol{a}; \boldsymbol{y})$ existe si y sólo si $f'_k(\boldsymbol{a}; \boldsymbol{y})$ existe para cada $k = 1, 2, \ldots, m$, en cuyo caso tenemos

$$f'(\boldsymbol{a}; \boldsymbol{y}) = (f'_1(\boldsymbol{a}; \boldsymbol{y}), \ldots, f'_m(\boldsymbol{a}; \boldsymbol{y})) = \sum_{k=1}^{m} f'_k(\boldsymbol{a}; \boldsymbol{y})\boldsymbol{e}_k,$$

donde $\boldsymbol{e}_k$ es el k-ésimo vector coordenado unidad.

Decimos que f es *diferenciable* en un punto interior $\boldsymbol{a}$ si existe una transformación lineal

$$T_{\boldsymbol{a}}: \mathbf{R}^n \to \mathbf{R}^m$$

tal que

(8.16) $$f(\boldsymbol{a} + \boldsymbol{v}) = f(\boldsymbol{a}) + T_{\boldsymbol{a}}(\boldsymbol{v}) + \|\boldsymbol{v}\|\, E(\boldsymbol{a}, \boldsymbol{v}),$$

donde $E(\boldsymbol{a}, \boldsymbol{v}) \to \boldsymbol{O}$ cuando $\boldsymbol{v} \to \boldsymbol{O}$. La fórmula de Taylor de primer orden (8.16) es válida para todo $\boldsymbol{v}$ tal que $\|\boldsymbol{v}\| < r$ para un cierto $r > 0$. El término $E(\boldsymbol{a}, \boldsymbol{v})$ es un vector de $\mathbf{R}^m$. La transformación lineal $T_{\boldsymbol{a}}$ se llama *diferencial total* o simplemente *diferencial* de f en $\boldsymbol{a}$.

Para los campos escalares se demostró que $T_{\boldsymbol{a}}(\boldsymbol{y})$ es el producto escalar del vector gradiente $\nabla f(\boldsymbol{a})$ por $\boldsymbol{y}$. Para los campos vectoriales demostraremos que $T_{\boldsymbol{a}}(\boldsymbol{y})$ es un vector cuyo componente k-ésimo es el producto escalar $\nabla f_k(\boldsymbol{a}) \cdot \boldsymbol{y}$.

TEOREMA 8.9. *Supongamos que f es diferenciable en $\boldsymbol{a}$ con diferencial $T_{\boldsymbol{a}}$. Existe entonces la derivada $f'(\boldsymbol{a}; \boldsymbol{y})$ para todo $\boldsymbol{a}$ de $\mathbf{R}^n$, y tenemos*

(8.17) $$T_{\boldsymbol{a}}(\boldsymbol{y}) = f'(\boldsymbol{a}; \boldsymbol{y}).$$

Además, si $f = (f_1, \ldots, f_m)$ y si $\boldsymbol{y} = (y_1, \ldots, y_n)$, tenemos

(8.18) $$T_{\boldsymbol{a}}(\boldsymbol{y}) = \sum_{k=1}^{m} \nabla f_k(\boldsymbol{a}) \cdot \boldsymbol{y}\, \boldsymbol{e}_k = (\nabla f_1(\boldsymbol{a}) \cdot \boldsymbol{y}, \ldots, \nabla f_m(\boldsymbol{a}) \cdot \boldsymbol{y}).$$

Demostración. Razonemos como en el caso escalar. Si $\boldsymbol{y} = \boldsymbol{O}$, entonces $f'(\boldsymbol{a}; \boldsymbol{y}) = \boldsymbol{O}$ y $T_{\boldsymbol{a}}(\boldsymbol{O}) = \boldsymbol{O}$. Por consiguiente podemos suponer que $\boldsymbol{y} \neq \boldsymbol{O}$. Tomando $\boldsymbol{v} = h\boldsymbol{y}$ en la fórmula de Taylor (8.16) tenemos

$$f(\boldsymbol{a} + h\boldsymbol{y}) - f(\boldsymbol{a}) = T_{\boldsymbol{a}}(h\boldsymbol{y}) + \|h\boldsymbol{y}\|\, E(\boldsymbol{a}, \boldsymbol{v}) = hT_{\boldsymbol{a}}(\boldsymbol{y}) + |h|\, \|\boldsymbol{y}\|\, E(\boldsymbol{a}, \boldsymbol{v}).$$

Dividiendo por h y haciendo que $h \to 0$ obtenemos (8.17).

Para demostrar (8.18) basta observar que

$$f'(\boldsymbol{a}; \boldsymbol{y}) = \sum_{k=1}^{m} f'_k(\boldsymbol{a}; \boldsymbol{y})\, \boldsymbol{e}_k = \sum_{k=1}^{m} \nabla f_k(\boldsymbol{a}) \cdot \boldsymbol{y}\, \boldsymbol{e}_k.$$

La ecuación (8.18) puede también escribirse en forma más sencilla como un producto matricial,

$$T_a(y) = Df(a)y,$$

siendo $Df(a)$ la matriz $m \times n$ cuya fila k-ésima es $\nabla f_k(a)$, e y una matriz columna $n \times 1$. La matriz $Df(a)$ se llama *matriz jacobiana* de f en a. Su elemento kj es la derivada parcial $D_j f_k(a)$. Así pues, tenemos

$$Df(a) = \begin{bmatrix} D_1 f_1(a) & D_2 f_1(a) & \cdots & D_n f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \cdots & D_n f_2(a) \\ \vdots & & & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \cdots & D_n f_m(a) \end{bmatrix}.$$

La matriz jacobiana $Df(a)$ está definida en cada punto en el que existan las mn derivadas parciales $D_j f_k(a)$.

La diferencial T_a se expresa también poniendo $f'(a)$. La derivada $f'(a)$ es una transformación lineal; la matriz jacobiana $Df(a)$ es una representación matricial de esa transformación.

La fórmula de Taylor de primer orden toma la forma

$$f(a + v) = f(a) + f'(a)(v) + \|v\| E(a, v), \tag{8.19}$$

donde $E(a, v) \to O$ cuando $v \to O$. Se parece a la fórmula de Taylor unidimensional. Para calcular los componentes del vector $f'(a)(v)$ podemos utilizar el producto matricial $Df(a)v$ o la fórmula (8.18) del teorema 8.9.

8.19 La diferenciabilidad implica la continuidad

TEOREMA 8.10. *Si un campo vectorial f es diferenciable en a, entonces f es continuo en a.*

Demostración. Como en el caso escalar, aplicamos la fórmula de Taylor para demostrar este teorema. Si hacemos que $v \to O$ en (8.19) el error $\|v\| E(a, v) \to O$. La parte lineal $f'(a)(v)$ tiende también a O debido a que las transformaciones lineales son continuas en O. Esto completa la demostración.

Al llegar aquí conviene deducir una desigualdad que se utilizará en la de-

mostración de la regla de la cadena en la próxima sección. La desigualdad se refiere a un campo vectorial f diferenciable en a; ella establece que

(8.20) $$\|f'(a)(v)\| \leq M_f(a)\,\|v\|, \qquad \text{donde} \quad M_f(a) = \sum_{k=1}^{m} \|\nabla f_k(a)\|.$$

Para demostrarla utilizamos la ecuación (8.18) junto a la desigualdad triangular de Cauchy-Schwarz obteniendo

$$\|f'(a)(v)\| = \left\| \sum_{k=1}^{m} \nabla f_k(a) \cdot v\, e_k \right\| \leq \sum_{k=1}^{m} |\nabla f_k(a) \cdot v| \leq \sum_{k=1}^{m} \|\nabla f_k(a)\|\, \|v\|.$$

8.20 La regla de la cadena para diferenciales de campos vectoriales

TEOREMA 8.11. REGLA DE LA CADENA. *Sean f y g dos campos vectoriales tales que la función compuesta $h = f \circ g$ esté definida en un entorno del punto a. Supongamos que g sea diferenciable en a, con diferencial $g'(a)$. Pongamos $b = g(a)$ y supongamos que f es diferenciable en b, con diferencial $f'(b)$. Entonces h es diferenciable en a, y la diferencial $h'(a)$ viene dada por*

$$h'(a) = f'(b) \circ g'(a),$$

que es la composición de las transformaciones lineales $f'(b)$ y $g'(a)$.

Demostración. Consideremos la diferencia $h(a + y) - h(a)$ para valores pequeños de $\|y\|$, y demostremos que se obtiene una fórmula de Taylor de primer orden. De la definición de h resulta

(8.21) $$h(a + y) - h(a) = f[g(a + y)] - f[g(a)] = f(b + v) - f(b),$$

siendo $v = g(a + y) - g(a)$. La fórmula de Taylor aplicada a $g(a + y)$ nos da

(8.22) $$v = g'(a)(y) + \|y\|\, E_g(a, y), \qquad \text{donde} \quad E_g(a, y) \to O \text{ cuando } y \to O.$$

La fórmula de Taylor relativa a $f(b + v)$ nos da

(8.23) $$f(b + v) - f(b) = f'(b)(v) + \|v\|\, E_f(b, v),$$

donde $E_f(b, v) \to O$ cuando $v \to O$. Aplicando (8.22) en (8.23) obtenemos

$$(8.24)\quad f(b + v) - f(b) = f'(b)g'(a)(y) + f'(b)(\|y\|\, E_g(a, y)) + \|v\|\, E_f(b, v)$$
$$= f'(b)g'(a)(y) + \|y\|\, E(a, y),$$

donde $E(a, O) = O$ y

$$(8.25)\qquad E(a, y) = f'(b)(E_g(a, y)) + \frac{\|v\|}{\|y\|}\, E_f(b, v) \qquad \text{si} \quad y \neq O.$$

Para completar la demostración necesitamos probar que $E(a, y) \to O$ cuando $y \to O$.

El primer término del segundo miembro de (8.25) tiende a O cuando $y \to O$ porque $E_g(a, y) \to O$ cuando $y \to O$ y las transformaciones lineales son continuas en O. .

En el segundo término del segundo miembro de (8.25) el factor $E_f(b, v) \to O$ porque $v \to O$ cuando $y \to O$. El cociente $\|v\|/\|y\|$ permanece acotado porque, según (8.22) y (8.20) tenemos

$$\|v\| \leq M_g(a)\, \|y\| + \|y\|\, \|E_g(a, y)\|.$$

Por consiguiente los dos términos del segundo miembro de (8.25) tienden a O cuando $y \to O$, así que $E(a, y) \to O$.

De este modo, de (8.24) y (8.21) obtenemos la fórmula de Taylor

$$h(a + y) - h(a) = f'(b)g'(a)(y) + \|y\|\, E(a, y),$$

donde $E(a, y) \to O$ cuando $y \to O$. Esto demuestra que h es diferenciable en a y que la diferencial $h'(a)$ es igual a la composición $f'(b) \circ g'(a)$.

8.21 Forma matricial de la regla de la cadena

Sea $h = f \circ g$, donde g es diferenciable en a y f diferenciable en $b = g(a)$. La regla de la cadena establece que

$$h'(a) = f'(b) \circ g'(a).$$

Podemos expresar la regla de la cadena en función de las matrices jacobianas $Dh(a)$, $Df(b)$, y $Dg(a)$ que representan las transformaciones lineales $h'(a)$, $f'(b)$, y $g'(a)$, respectivamente. Puesto que la composición de transformaciones lineales corresponde a la multiplicación de sus matrices, obtenemos

$$(8.26)\qquad Dh(a) = Df(b)\, Dg(a), \qquad \text{donde} \quad b = g(a).$$

Ésta es la llamada *forma matricial de la regla de la cadena.* También puede escribirse como un conjunto de ecuaciones escalares expresando cada matriz en función de sus elementos.

Supongamos que $\boldsymbol{a} \in \mathbf{R}^p$, $\boldsymbol{b} = \boldsymbol{g}(\boldsymbol{a}) \in \mathbf{R}^n$, y $\boldsymbol{f}(\boldsymbol{b}) \in \mathbf{R}^m$. Entonces $\boldsymbol{h}(\boldsymbol{a}) \in \mathbf{R}^m$ y podemos escribir

$$\boldsymbol{g} = (g_1, \ldots, g_n), \qquad \boldsymbol{f} = (f_1, \ldots, f_m), \qquad \boldsymbol{h} = (h_1, \ldots, h_m).$$

La matriz $D\boldsymbol{h}(\boldsymbol{a})$ es $m \times p$, la matriz $D\boldsymbol{f}(\boldsymbol{b})$ es $m \times n$, y $D\boldsymbol{g}(\boldsymbol{a})$ es una matriz $n \times p$, y vienen dadas por

$$D\boldsymbol{h}(\boldsymbol{a}) = [D_j h_i(\boldsymbol{a})]_{i,j=1}^{m,p}, \qquad D\boldsymbol{f}(\boldsymbol{b}) = [D_k f_i(\boldsymbol{b})]_{i,k=1}^{m,n}, \qquad D\boldsymbol{g}(\boldsymbol{a}) = [D_j g_k(\boldsymbol{a})]_{k,j=1}^{n,p}.$$

La ecuación matricial (8.26) es equivalente a mp ecuaciones escalares,

$$D_j h_i(\boldsymbol{a}) = \sum_{k=1}^{n} D_k f_i(\boldsymbol{b}) D_j g_k(\boldsymbol{a}), \quad \text{para} \quad i = 1, 2, \ldots, m \qquad \text{y} \qquad j = 1, 2, \ldots, p.$$

Estas ecuaciones expresan las derivadas parciales de los componentes de $\boldsymbol{h}$ en función de las derivadas parciales de los componentes de $\boldsymbol{f}$ y $\boldsymbol{g}$.

EJEMPLO 1. *Extensión de la regla de la cadena para campos escalares.* Supongamos que f es un campo escalar ($m = 1$). Entonces h también lo es y existen p ecuaciones en la regla de la cadena, una para cada una de las derivadas parciales de h:

$$D_j h(\boldsymbol{a}) = \sum_{k=1}^{n} D_k f(\boldsymbol{b}) D_j g_k(\boldsymbol{a}), \quad \text{para} \ j = 1, 2, \ldots, p.$$

El caso particular $p = 1$ ya se consideró en la sección 8.15. Entonces se tiene la única ecuación,

$$h'(\boldsymbol{a}) = \sum_{k=1}^{n} D_k f(\boldsymbol{b}) g_k'(\boldsymbol{a}).$$

Consideremos ahora $p = 2$ y $n = 2$. Pongamos $\boldsymbol{a} = (s, t)$ y $\boldsymbol{b} = (x, y)$. Los componentes x e y están ligados a s y t por las ecuaciones

$$x = g_1(s, t), \qquad y = g_2(s, t).$$

La regla de la cadena nos da un par de ecuaciones para las derivadas parciales de h:

$$D_1h(s, t) = D_1f(x, y)\, D_1g_1(s, t) + D_2f(x, y)\, D_1g_2(s, t),$$
$$D_2h(s, t) = D_1f(x, y)\, D_2g_1(s, t) + D_2f(x, y)\, D_2g_2(s, t).$$

Empleando el signo ∂, también se pone este par de ecuaciones en la forma

(8.27) $$\frac{\partial h}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}, \qquad \frac{\partial h}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}.$$

EJEMPLO 2. *Coordenadas polares.* La temperatura de una placa delgada se representa por un campo escalar f, siendo $f(x, y)$ la temperatura en (x, y). Introduciendo las coordenadas polares $x = r \cos \theta$, $y = r \operatorname{sen} \theta$, la temperatura se convierte en una función de r y θ determinada por medio de la ecuación

$$\varphi(r, \theta) = f(r \cos \theta, r \operatorname{sen} \theta).$$

Expresar las derivadas parciales $\partial\varphi/\partial r$ y $\partial\varphi/\partial\theta$ en función de las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$.

Solución. Utilicemos la regla de la cadena en la forma (8.27), poniendo (r, θ) en lugar de (s, t), y φ en lugar de h. Las ecuaciones

$$x = r \cos \theta, \qquad y = r \operatorname{sen} \theta$$

nos dan

$$\frac{\partial x}{\partial r} = \cos \theta, \qquad \frac{\partial y}{\partial r} = \operatorname{sen} \theta, \qquad \frac{\partial x}{\partial \theta} = -r \operatorname{sen} \theta, \qquad \frac{\partial y}{\partial \theta} = r \cos \theta.$$

Sustituyendo esas fórmulas en (8.27) obtenemos

(8.28) $$\frac{\partial \varphi}{\partial r} = \frac{\partial f}{\partial x}\cos \theta + \frac{\partial f}{\partial y}\operatorname{sen}\theta, \qquad \frac{\partial \varphi}{\partial \theta} = -r\frac{\partial f}{\partial x}\operatorname{sen}\theta + r\frac{\partial f}{\partial y}\cos \theta.$$

Éstas son las fórmulas pedidas correspondientes a $\partial\varphi/\partial r$ y $\partial\varphi/\partial\theta$.

EJEMPLO 3. *Derivadas parciales de segundo orden.* Continuando el ejemplo 2, expresar la derivada parcial de segundo orden $\partial^2\varphi/\partial\theta^2$ en función de las derivadas parciales de f.

Solución. Comencemos con la fórmula que da $\partial\varphi/\partial\theta$ en (8.28) y derivemos

respecto a θ, considerando r como una constante. En el segundo miembro hay dos términos, cada uno de los cuales debe derivarse como producto. Obtenemos así

$$\begin{aligned}\frac{\partial^2\varphi}{\partial\theta^2} &= -r\frac{\partial f}{\partial x}\frac{\partial(\operatorname{sen}\theta)}{\partial\theta} - r\operatorname{sen}\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) + r\frac{\partial f}{\partial y}\frac{\partial(\cos\theta)}{\partial\theta} + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right)\\ &= -r\cos\theta\frac{\partial f}{\partial x} - r\operatorname{sen}\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) - r\operatorname{sen}\theta\frac{\partial f}{\partial y} + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right).\end{aligned} \tag{8.29}$$

Para calcular las derivadas de $\partial f/\partial x$ y $\partial f/\partial y$ respecto a θ debe tenerse en cuenta que como funciones de r y θ, $\partial f/\partial x$ y $\partial f/\partial y$ son *funciones compuestas* dadas por

$$\frac{\partial f}{\partial x} = D_1 f(r\cos\theta, r\operatorname{sen}\theta) \qquad \text{y} \qquad \frac{\partial f}{\partial y} = D_2 f(r\cos\theta, r\operatorname{sen}\theta).$$

Por consiguiente, sus derivadas respecto a θ tienen que determinarse con la regla de la cadena. Apliquemos otra vez (8.27), reemplazando f por $D_1 f$, con lo que se obtiene

$$\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial(D_1 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_1 f)}{\partial y}\frac{\partial y}{\partial\theta} = \frac{\partial^2 f}{\partial x^2}(-r\operatorname{sen}\theta) + \frac{\partial^2 f}{\partial y\,\partial x}(r\cos\theta).$$

Del mismo modo, aplicando (8.27) reemplazando f por $D_2 f$, encontramos

$$\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial(D_2 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_2 f)}{\partial y}\frac{\partial y}{\partial\theta} = \frac{\partial^2 f}{\partial x\,\partial y}(-r\operatorname{sen}\theta) + \frac{\partial^2 f}{\partial y^2}(r\cos\theta).$$

Cuando estas fórmulas se aplican en (8.29) obtenemos

$$\begin{aligned}\frac{\partial^2\varphi}{\partial\theta^2} = -r\cos\theta\frac{\partial f}{\partial x} &+ r^2\operatorname{sen}^2\theta\frac{\partial^2 f}{\partial x^2} - r^2\operatorname{sen}\theta\cos\theta\frac{\partial^2 f}{\partial y\,\partial x}\\ &- r\operatorname{sen}\theta\frac{\partial f}{\partial y} - r^2\operatorname{sen}\theta\cos\theta\frac{\partial^2 f}{\partial x\,\partial y} + r^2\cos^2\theta\frac{\partial^2 f}{\partial y^2}.\end{aligned}$$

Ésta es la fórmula que deseábamos para $\partial^2\varphi/\partial\theta^2$. Fórmulas análogas para las derivadas parciales segundas $\partial^2\varphi/\partial r^2$, $\partial^2\varphi/(\partial r\,\partial\theta)$, y $\partial^2\varphi/(\partial\theta\,\partial r)$ se proponen en el ejercicio 5 de la próxima sección.

8.22 Ejercicios

En estos ejercicios puede suponerse la diferenciabilidad de todas las funciones que se consideran.

1. La sustitución $t = g(x, y)$ convierte $F(t)$ en $f(x, y)$, siendo $f(x, y) = F[g(x, y)]$.
a) Demostrar que

$$\frac{\partial f}{\partial x} = F'[g(x, y)]\frac{\partial g}{\partial x} \quad \text{y} \quad \frac{\partial f}{\partial y} = F'[g(x, y)]\frac{\partial g}{\partial y}.$$

b) Considérese el caso particular $F(t) = e^{\operatorname{sen} t}$, $g(x, y) = \cos(x^2 + y^2)$. Calcular $\partial f/\partial x$ y $\partial f/\partial y$ utilizando las fórmulas del apartado a). Comprobar el resultado, determinando $f(x, y)$ explícitamente en función de x e y, y calculando directamente $\partial f/\partial x$ y $\partial f/\partial y$ a partir de f.
2. La sustitución $u = (x - y)/2$, $v = (x + y)/2$ cambia $f(u, v)$ en $F(x, y)$. Aplicar en forma adecuada la regla de la cadena para expresar las derivadas parciales $\partial F/\partial x$ y $\partial F/\partial y$ en función de las derivadas parciales $\partial f/\partial u$ y $\partial f/\partial v$.
3. Las ecuaciones $u = f(x, y)$, $x = X(s, t)$, $y = Y(s, t)$ definen u como función de s y t, $u = F(s, t)$.
a) Aplicar una forma adecuada de la regla de la cadena para expresar las derivadas parciales $\partial F/\partial s$ y $\partial F/\partial t$ en función de $\partial f/\partial x$, $\partial f/\partial y$, $\partial X/\partial s$, $\partial X/\partial t$, $\partial Y/\partial s$, $\partial Y/\partial t$.
b) Si $\partial^2 f/(\partial x\, \partial y) = \partial^2 f/(\partial y\, \partial x)$, demostrar que

$$\frac{\partial^2 F}{\partial s^2} = \frac{\partial f}{\partial x}\frac{\partial^2 X}{\partial s^2} + \frac{\partial^2 f}{\partial x^2}\left(\frac{\partial X}{\partial s}\right)^2 + 2\frac{\partial X}{\partial s}\frac{\partial Y}{\partial s}\frac{\partial^2 f}{\partial x\, \partial y} + \frac{\partial f}{\partial y}\frac{\partial^2 Y}{\partial s^2} + \frac{\partial^2 f}{\partial y^2}\left(\frac{\partial Y}{\partial s}\right)^2.$$

c) Encontrar fórmulas parecidas para las derivadas parciales $\partial^2 F/(\partial s\, \partial t)$ y $\partial^2 F/\partial t^2$.
4. Resolver el ejercicio 3 en cada uno de los siguientes casos particulares:
a) $X(s, t) = s + t$, $Y(s, t) = st$.
b) $X(s, t) = st$, $Y(s, t) = s/t$.
c) $X(s, t) = (s - t)/2$, $Y(s, t) = (s + t)/2$.
5. La introducción de las coordenadas polares cambia $f(x, y)$ en $\varphi(r, \theta)$, donde $x = r\cos\theta$ e $y = r \operatorname{sen}\theta$. Expresar las derivadas parciales de segundo orden $\partial^2\varphi/\partial r^2$, $\partial^2\varphi/(\partial r\, \partial\theta)$ y $\partial^2\varphi/(\partial\theta\, \partial r)$ en función de las derivadas parciales de f. Pueden utilizarse las fórmulas deducidas en el ejemplo 2 de la sección 8.21.
6. Las ecuaciones $u = f(x, y, z)$, $x = X(r, s, t)$, $y = Y(r, s, t)$ y $z = Z(r, s, t)$, definen u como función de r, s y t, sea ésta $u = F(r, s, t)$. Aplicar la forma adecuada de la regla de la cadena para expresar las derivadas parciales $\partial F/\partial r$, $\partial F/\partial s$, y $\partial F/\partial t$ en función de las derivadas parciales de f, X, Y y Z.
7. Resolver el ejercicio 6 en cada uno de los casos particulares siguientes:
a) $X(r, s, t) = r + s + t$, $Y(r, s, t) = r - 2s + 3t$, $Z(r, s, t) = 2r + s - t$.
b) $X(r, s, t) = r^2 + s^2 + t^2$, $Y(r, s, t) = r^2 - s^2 - t^2$, $Z(r, s, t) = r^2 - s^2 + t^2$.
8. Las ecuaciones $u = f(x, y, z)$, $x = X(s, t)$, $y = Y(s, t)$, $z = Z(s, t)$ define u como función de s y t, sea ésta $u = F(s, t)$. Aplicando una forma adecuada de la regla de la cadena expresar las derivadas parciales $\partial F/\partial s$ y $\partial F/\partial t$ en función de las derivadas parciales de f, X, Y y Z.

9. Resolver el ejercicio 8 en cada uno de los casos particulares siguientes:
   a) $X(s,t) = s^2 + t^2, \quad Y(s,t) = s^2 - t^2, \quad Z(s,t) = 2st.$
   b) $X(s,t) = s + t, \quad Y(s,t) = s - t, \quad Z(s,t) = st.$
10. Las ecuaciones $u = f(x,y)$, $x = X(r,s,t)$, $y = Y(r,s,t)$ definen u como función de r, s y t, sea ésta $u = F(r,s,t)$. Aplicar una forma adecuada de la regla de la cadena y expresar las derivadas parciales $\partial F/\partial r$, $\partial F/\partial s$ y $\partial F/\partial t$ en función de las derivadas parciales de f, X e Y.
11. Resolver el ejercicio 10 en cada uno de los casos particulares siguientes:
   a) $X(r,s,t) = r + s, \quad Y(r,s,t) = t.$
   b) $X(r,s,t) = r + s + t, \quad Y(r,s,t) = r^2 + s^2 + t^2.$
   c) $X(r,s,t) = r/s, \quad Y(r,s,t) = s/t.$
12. Sea $h(\boldsymbol{x}) = f[\boldsymbol{g}(\boldsymbol{x})]$, donde $\boldsymbol{g} = (g_1, \ldots, g_n)$ es un campo vectorial diferenciable en $\boldsymbol{a}$, y f un campo escalar diferenciable en $\boldsymbol{b} = \boldsymbol{g}(\boldsymbol{a})$. Utilizar la regla de la cadena para demostrar que el gradiente de h puede expresarse como combinación lineal de los vectores gradientes de los componentes de $\boldsymbol{g}$, así:

$$\nabla h(\boldsymbol{a}) = \sum_{k=1}^{n} D_k f(\boldsymbol{b}) \nabla g_k(\boldsymbol{a}).$$

13. a) Si $\boldsymbol{f}(x,y,z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ demostrar que la matriz jacobiana $D\boldsymbol{f}(x,y,z)$ es la matriz identidad de orden 3.
   b) Hallar todos los campos vectoriales diferenciables $\boldsymbol{f}: \mathbf{R}^3 \to \mathbf{R}^3$ para los que la matriz jacobiana $D\boldsymbol{f}(x,y,z)$ es la matriz identidad de orden 3.
   c) Hallar todos los campos vectoriales diferenciables $\boldsymbol{f}: \mathbf{R}^3 \to \mathbf{R}^3$ para los que la matriz jacobiana es una matriz diagonal de la forma diag $(p(x), q(y), r(z))$, en la que p, q y r son funciones continuas dadas.
14. Sean $\boldsymbol{f}: \mathbf{R}^2 \to \mathbf{R}^2$ y $\boldsymbol{g}: \mathbf{R}^3 \to \mathbf{R}^2$ dos campos vectoriales definidos del siguiente modo:

$$\boldsymbol{f}(x,y) = e^{x+2y}\boldsymbol{i} + \operatorname{sen}(y + 2x)\boldsymbol{j},$$

$$\boldsymbol{g}(u,v,w) = (u + 2v^2 + 3w^3)\boldsymbol{i} + (2v - u^2)\boldsymbol{j}.$$

   a) Calcular cada una de las matrices jacobianas $D\boldsymbol{f}(x,y)$ y $D\boldsymbol{g}(u,v,w)$.
   b) Calcular la función compuesta $\boldsymbol{h}(u,v,w) = \boldsymbol{f}[\boldsymbol{g}(u,v,w)]$.
   c) Calcular la matriz jacobiana $D\boldsymbol{h}(1,-1,1)$.
15. Sean $\boldsymbol{f}: \mathbf{R}^3 \to \mathbf{R}^2$ y $\boldsymbol{g}: \mathbf{R}^3 \to \mathbf{R}^3$ dos campos vectoriales definidos como sigue:

$$\boldsymbol{f}(x,y,z) = (x^2 + y + z)\boldsymbol{i} + (2x + y + z^2)\boldsymbol{j},$$

$$\boldsymbol{g}(u,v,w) = uv^2w^2\boldsymbol{i} + w^2 \operatorname{sen} v\,\boldsymbol{j} + u^2e^v\boldsymbol{k}.$$

   a) Calcular cada una de las matrices jacobianas $D\boldsymbol{f}(x,y,z)$ y $D\boldsymbol{g}(u,v,w)$.
   b) Calcular la función compuesta $\boldsymbol{h}(u,v,w) = \boldsymbol{f}[\boldsymbol{g}(u,v,w)]$.
   c) Calcular la matriz jacobiana $D\boldsymbol{h}(u,0,w)$.

* 8.23 Condiciones suficientes para la igualdad de las derivadas parciales mixtas

Si f es una función real de dos variables, las dos derivadas parciales mixtas $D_{1,2}f$ y $D_{2,1}f$ no son necesariamente iguales. Con la notación $D_{1,2}f$ queremos

significar $D_1(D_2 f) = \partial^2 f/(\partial x\, \partial y)$ y con $D_{2,1} f$ significar $D_2(D_1 f) = \partial^2 f/(\partial y\, \partial x)$. Por ejemplo, si f está definida por las ecuaciones

$$f(x, y) = xy\frac{x^2 - y^2}{x^2 + y^2} \quad \text{para} \quad (x, y) \neq (0, 0), \qquad f(0, 0) = 0,$$

es sencillo demostrar que $D_{2,1}f(0, 0) = -1$ y $D_{1,2}f(0, 0) = 1$. Esto puede verse así:

La definición de $D_{2,1}f(0, 0)$ establece que

$$(8.30) \qquad D_{2,1}f(0, 0) = \lim_{k\to 0} \frac{D_1 f(0, k) - D_1 f(0, 0)}{k}.$$

Ahora bien

$$D_1 f(0, 0) = \lim_{h\to 0} \frac{f(h, 0) - f(0, 0)}{h} = 0$$

y, si $(x,y) \neq (0, 0)$, encontramos

$$D_1 f(x, y) = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2}.$$

Por consiguiente, si $k \neq 0$ tenemos $D_1 f(0, k) = -k^5/k^4 = -k$ y por tanto

$$\frac{D_1 f(0, k) - D_1 f(0, 0)}{k} = -1.$$

Aplicando este resultado en (8.30) encontramos que $D_{2,1}f(0, 0) = -1$. Con razonamiento parecido demostramos que $D_{1,2}f(0,0) = 1$, y por tanto $D_{2,1}f(0, 0) \neq D_{1,2}f(0, 0)$.

En el ejemplo que acabamos de considerar las dos derivadas parciales mixtas $D_{1,2}f$ y $D_{2,1}f$ son ambas no continuas en el origen. Puede demostrarse que las dos parciales mixtas *son* iguales en un punto (a, b) si por lo menos una de ellas es continua en el punto. Demostraremos primero que son iguales si *ambas* son continuas. Con mayor precisión, tenemos el teorema siguiente.

TEOREMA 8.12. CONDICIÓN SUFICIENTE PARA LA IGUALDAD DE LAS DERIVADAS PARCIALES MIXTAS. *Si f es un campo escalar tal que las derivadas parciales $D_1 f$, $D_2 f$, $D_{1,2}f$ y $D_{2,1}f$ existen en un conjunto abierto S. Si (a, b) es un punto de S, y en tal punto $D_{1,2}f$ y $D_{2,1}f$ son continuas, entonces tenemos*

$$(8.31) \qquad D_{1,2}f(a, b) = D_{2,1}f(a, b).$$

Demostración. Elijamos h y k no nulos tales que el rectángulo $R(h, k)$

con vértices (a, b), $(a + h, b)$, $(a + h, b + k)$ y $(a, b + k)$ esté situado en S. (Véase figura 8.9.)

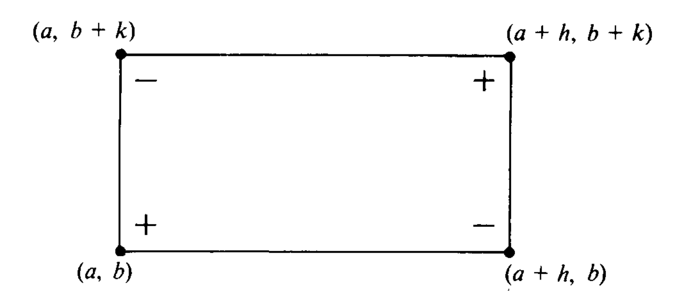


FIGURA 8.9 $\nabla(h,k)$ *es una combinación de los valores de f en los vértices.*

Consideremos la expresión

$$\Delta(h, k) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b).$$

Es ésta una combinación de los valores de f en los vértices de $R(h, k)$, tomados con los signos algebraicos indicados en la figura 8.9. Expresaremos $\Delta(h, k)$ en función de $D_{2,1}f$ y también de $D_{1,2}f$.

Consideremos una nueva función G de una variable definida por la ecuación

$$G(x) = f(x, b + k) - f(x, b)$$

para todo x comprendido entre a y $a + h$. (Geométricamente, consideramos los valores de f en aquellos puntos en los que una recta vertical corta los lados horizontales de $R(h, k$.) Tenemos entonces

$$\textbf{(8.32)} \qquad \Delta(h, k) = G(a + h) - G(a).$$

Aplicando el teorema del valor medio uni-dimensional al segundo miembro de (8.3) obtenemos $G(a + h) - G(a) = hG'(x_1)$, siendo x_1 un punto situado entre a y $a + h$. Puesto que $G'(x) = D_1f(x, b + k) - D_1f(x, b)$, la ecuación (8.32) se transforma en

$$(8.33) \qquad \Delta(h, k) = h[D_1f(x_1, b + k) - D_1f(x_1, b)].$$

Aplicando el teorema del valor medio al segundo miembro de (8.33) obtenemos

$$(8.34) \qquad \Delta(h, k) = hkD_{2,1}f(x_1, y_1),$$

siendo y_1 un punto situado entre b y $b + k$. El punto (x_1, y_1) pertenece al rectángulo $R(h, k)$.

Aplicando el mismo procedimiento a la función $H(y) = f(a + h, y) - f(a, y)$ encontramos una segunda expresión para $\Delta(h, k)$, o sea,

$$\Delta(h, k) = hkD_{1,2}f(x_2, y_2), \tag{8.35}$$

donde (x_2, y_2) pertenece también a $R(h, k)$. Igualando las dos expresiones de $\Delta(h, k)$ y suprimiendo hk obtenemos

$$D_{1,2}f(x_1, y_1) = D_{2,1}f(x_2, y_2).$$

Hagamos ahora que $(h, k) \to (0, 0)$ y teniendo en cuenta la continuidad de $D_{1,2}f$ y $D_{2,1}f$ en el punto (a, b) obtenemos (8.31).

El razonamiento anterior puede modificarse para demostrar una versión más fuerte del teorema 8.12.

TEOREMA 8.13. *Si f es un campo escalar para el cual existen las derivadas parciales D_1f, D_2f y $D_{2,1}f$ en un conjunto abierto S que contenga (a, b), y si además $D_{2,1}f$ es continua en S, entonces existe la derivada $D_{1,2}f(a, b)$ y tenemos*

$$D_{1,2}f(a, b) = D_{2,1}f(a, b).$$

Demostración. Definamos $\Delta(h, k)$ como en la demostración del teorema 8.12. La parte de la demostración que lleva a la ecuación (8.34) es válida, dándonos

$$\frac{\Delta(h, k)}{hk} = D_{2,1}f(x_1, y_1) \tag{8.36}$$

para un cierto (x_1, y_1) del rectángulo $R(h, k)$. El resto de la demostración no es aplicable ya que precisa de la existencia de la derivada $D_{1,2}f(a, b)$, que es justamente lo que deseamos demostrar.

La definición de $D_{1,2}f(a, b)$ establece que

$$D_{1,2}f(a, b) = \lim_{h\to 0} \frac{D_2f(a + h, b) - D_2f(a, b)}{h}. \tag{8.37}$$

Vamos a demostrar que este límite existe y que tiene el valor $D_{2,1}f(a, b)$. A partir de la definición de D_2f tenemos

$$D_2f(a, b) = \lim_{k\to 0} \frac{f(a, b + k) - f(a, b)}{k}$$

y

$$D_2f(a + h, b) = \lim_{k\to 0} \frac{f(a + h, b + k) - f(a + h, b)}{k}.$$

Por tanto el cociente de diferencias (8.37) puede escribirse del siguiente modo

$$\frac{D_2f(a + h, b) - D_2f(a, b)}{h} = \lim_{k\to 0} \frac{\Delta(h, k)}{hk}.$$

Teniendo en cuenta (8.36) podemos ponerlo en la forma

(8.38) $$\frac{D_2f(a + h, b) - D_2f(a, b)}{h} = \lim_{k\to 0} D_{2,1}f(x_1, y_1).$$

Para completar la demostración tenemos que probar que

(8.39) $$\lim_{h\to 0} \left[\lim_{k\to 0} D_{2,1}f(x_1, y_1)\right] = D_{2,1}f(a, b).$$

Cuando $k \to 0$, el punto $y_1 \to b$, pero es desconocido el comportamiento de x_1 como función de k. Si suponemos que x_1 se aproxima a algún límite, sea $\bar{x}$, cuando $k \to 0$, entonces por la continuidad de $D_{2,1}f$ deducimos que

$$\lim_{k\to 0} D_{2,1}f(x_1, y_1) = D_{2,1}f(\bar{x}, b).$$

Puesto que el límite $\bar{x}$ estaría en el intervalo $a \leq \bar{x} \leq a+h$, podemos suponer que $h \to 0$ y deducir (8.39). No obstante, por el hecho de que $\bar{x}$ depende de k de forma desconocida, se hace necesario un argumento algo más sólido.

En virtud de la ecuación (8.38) sabemos que existe el siguiente límite:

$$\lim_{k\to 0} D_{2,1}f(x_1, y_1).$$

Designemos este límite con $F(h)$. Para completar la demostración hay que probar que

$$\lim_{h\to 0} F(h) = D_{2,1}f(a, b).$$

A tal fin apelamos a la definición de continuidad de $D_{2,1}f$ en (a, b).

Sea ϵ un número positivo dado. La continuidad de $D_{2,1}f$ en (a, b) significa que existe un disco abierto N con centro en (a, b) y radio δ, por ejemplo, tal que

$$(8.40) \qquad |D_{2,1}f(x, y) - D_{2,1}f(a, b)| < \frac{\epsilon}{2} \quad \text{siempre que} \quad (x, y) \in N.$$

Si elegimos h y k de manera que $|h| < \delta/2$ y $|k| < \delta/2$, todo el rectángulo dibujado en la figura 8.9 estará contenido en el entorno N y, en particular, el punto (x_1, y_1) estará en N. Por consiguiente (8.40) es válida cuando $(x, y) = (x_1, y_1)$ y podemos escribir

$$(8.41) \qquad 0 \leq |D_{2,1}f(x_1, y_1) - D_{2,1}f(a, b)| < \frac{\epsilon}{2}.$$

Mantengamos ahora fijo h y hagamos que $k \to 0$. El término $D_{2,1}f(x_1, y_1)$ tiende a $F(h)$ y los otros términos de (8.41) son independientes de k. Tenemos por tanto

$$0 \leq |F(h) - D_{2,1}f(a, b)| \leq \frac{\epsilon}{2} < \epsilon,$$

con tal que $0 < |h| < \delta/2$. Pero éste es precisamente el significado de la proposición

$$\lim_{h\to 0} F(h) = D_{2,1}f(a, b)$$

y, como ya dijimos, esto completa la demostración.

Nota: Deberá observarse que el teorema es también válido si en él enunciado intercambiamos $D_{1,2}f$ y $D_{2,1}f$.

8.24 Ejercicios varios

1. Hallar un campo escalar f que satisfaga las dos condiciones siguientes·

a) Existen y son nulas las derivadas parciales $D_1f(0,0)$ y $D_2f(0,0)$.
b) Existe la derivada direccional en el origen en la dirección del vector $\boldsymbol{i}+\boldsymbol{j}$ y tiene el valor 3. Explicar por qué tal función f no puede ser diferenciable en (0,0).

2. Sea f una función definida como sigue:

$$f(x,y) = y\frac{x^2-y^2}{x^2+y^2} \quad \text{si} \quad (x,y) \neq (0,0), \qquad f(0,0)=0.$$

Calcular, cuando existan, las derivadas parciales siguientes: $D_1f(0,0)$, $D_2f(0,0)$, $D_{2,1}f(0,0)$, $D_{1,2}f(0,0)$.

3. Sea $f(x,y) = \dfrac{xy^3}{x^3+y^6}$ si $(x,y) \neq (0,0)$, y definamos $f(0,0)=0$.

a) Demostrar que la derivada $f'(\boldsymbol{O};\boldsymbol{a})$ existe para todo vector $\boldsymbol{a}$ y calcular su valor en función de los componentes de $\boldsymbol{a}$.
b) Determinar si f es o no continua en el origen.

4. Definamos $f(x,y) = \int_0^{\sqrt{xy}} e^{-t^2}\,dt$ para $x>0$, $y>0$. Calcular $\partial f/\partial x$ en función de x e y.

5. Supongamos que las ecuaciones $u=f(x,y)$, $x=X(t)$, $y=Y(t)$ definen u como función de t, $u=F(t)$. Calcular la derivada tercera $F'''(t)$ en función de las derivadas de $f, X,$ e Y.

6. El cambio de variables $x=u+v$, $y=uv^2$ transforma $f(x,y)$ en $g(u,v)$. Calcular el valor de $\partial^2 g/(\partial v\,\partial u)$ en el punto en el que $u=1$, $v=1$, sabiendo que

$$\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x} = 1$$

en dicho punto.

7. El cambio de variables $x=uv$, $y=\frac{1}{2}(u^2-v^2)$ transforma $f(x,y)$ en $g(u,v)$.
a) Calcular $\partial g/\partial u$, $\partial g/\partial v$ y $\partial^2 g/(\partial u\,\partial v)$en función de las derivadas parciales de f. (Puede suponerse la igualdad de las parciales mixtas.)
b) Si $\|\nabla f(x,y)\|^2 = 2$ para todo x e y, determinar las constantes a y b tales que

$$a\left(\frac{\partial g}{\partial u}\right)^2 - b\left(\frac{\partial g}{\partial v}\right)^2 = u^2+v^2.$$

8. Dos funciones F y G de una variable y una función z de dos variables están ligadas por la ecuación

$$[F(x)+G(y)]^2\, e^{z(x,y)} = 2F'(x)G'(y)$$

con tal que $F(x)+G(y)\neq 0$. Demostrar que la derivada parcial mixta $D_{2,1}z(x,y)$ nunca es cero. (Puede suponerse la existencia y continuidad de todas las derivadas que aparezcan.)

9. Un campo escalar f es acotado y continuo en un rectángulo $R=[a,b]\times[c,d]$. Se define en R un nuevo campo escalar g del modo siguiente:

$$g(u,v) = \int_c^v \left[\int_a^u f(x,y)\,dx\right] dy.$$

a) Puede demostrarse que para cada u fija en $[a, b]$ la función A definida en $[c, d]$ mediante la ecuación $A(y) = \int_a^u f(x, y)\,dx$ es continua en $[c, d]$. Utilizar este resultado para demostrar que $\partial g/\partial v$ existe y es continua en el rectángulo abierto $S = (a, b)\times(c, d)$ (el interior de R).
b) Supóngase que

$$\int_c^v \left[\int_a^u f(x, y)\,dx\right] dy = \int_a^u \left[\int_c^v f(x, y)\,dy\right] dx$$

para todo (u, v) de R. Demostrar que g es diferenciable en S y que las derivadas parciales mixtas $D_{1,2}g(u, v)$ y $D_{2,1}g(u, v)$ existen y son iguales a $f(u, v)$ en cada punto de S.

10. En relación con el ejercicio 9. Supóngase que u y v se expresan paramétricamente del siguiente modo: $u = A(t)$, $v = B(t)$; y sea $\varphi(t) = g[A(t), B(t)]$.
a) Determinar $\varphi'(t)$ en función de f, A' y B'.
b) Calcular $\varphi'(t)$ en función de t cuando $f(x, y) = e^{x+y}$ y $A(t) = B(t) = t^2$. (Supóngase que R está situado en el primer cuadrante.)
11. Si $f(x, y, z) = (\boldsymbol{r} \times \boldsymbol{A}) \cdot (\boldsymbol{r} \times \boldsymbol{B})$, siendo $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ y $\boldsymbol{A}$ y $\boldsymbol{B}$ son vectores constantes, demostrar que $\nabla f(x, y, z) = \boldsymbol{B} \times (\boldsymbol{r} \times \boldsymbol{A}) + \boldsymbol{A} \times (\boldsymbol{r} \times \boldsymbol{B})$.
12. Sea $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ y pongamos $r = \|\boldsymbol{r}\|$. Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son vectores constantes, demostrar que:

a) $\boldsymbol{A} \cdot \nabla\left(\dfrac{1}{r}\right) = -\dfrac{\boldsymbol{A} \cdot \boldsymbol{r}}{r^3}$.

b) $\boldsymbol{B} \cdot \nabla\left(\boldsymbol{A} \cdot \nabla\left(\dfrac{1}{r}\right)\right) = \dfrac{3\boldsymbol{A} \cdot \boldsymbol{r}\,\boldsymbol{B} \cdot \boldsymbol{r}}{r^5} - \dfrac{\boldsymbol{A} \cdot \boldsymbol{B}}{r^3}$

13. Hallar el conjunto de todos los puntos (a, b, c), en el espacio de 3 dimensiones, en los cuales las dos esferas $(x - a)^2 + (y - b)^2 + (z - c)^2 = 1$ y $x^2 + y^2 + z^2 = 1$ se cortan ortogonalmente. (Sus planos tangentes deberán ser perpendiculares en cada punto de intersección.)
14. Un cilindro cuya ecuación es $y = f(x)$ es tangente a la superficie $z^2 + 2xz + y = 0$ en todos los puntos comunes a las dos superficies. Hallar $f(x)$.

9

APLICACIONES DE CÁLCULO DIFERENCIAL

9.1 Ecuaciones diferenciales en derivadas parciales

Los teoremas de Cálculo diferencial desarrollados en el capítulo 8 tienen gran número de aplicaciones. Este capítulo muestra su utilización en algunos ejemplos relativos a las ecuaciones diferenciales en derivadas parciales, a las funciones implícitas y a problemas de extremos. Comenzamos con algunas observaciones elementales referentes a las ecuaciones diferenciales en derivadas parciales.

Una ecuación que relaciona un campo escalar f y sus derivadas parciales se llama *ecuación diferencial en derivadas parciales.* Dos ejemplos sencillos en los que f es una función de dos variables son la ecuación de primer orden

(9.1)
$$\frac{\partial f(x, y)}{\partial x} = 0,$$

y la de segundo orden

(9.2)
$$\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} = 0.$$

Cada una de ellas es una ecuación diferencial en derivadas parciales *lineal* homogénea. Esto es, cada una tiene la forma $L(f) = 0$, siendo L un operador diferencial lineal que contiene una o más derivadas parciales. La ecuación (9.2) se llama *ecuación de Laplace* bi-dimensional.

Parte de la teoría de las ecuaciones diferenciales lineales ordinarias puede extenderse a las ecuaciones diferenciales en derivadas parciales. Por ejemplo, es fácil comprobar que para cada una de las ecuaciones (9.1) y (9.2) el conjunto de soluciones es un espacio lineal. Sin embargo, hay una diferencia importante

entre las ecuaciones diferenciales lineales en derivadas parciales y las lineales ordinarias, que debe ser mencionada desde el principio. Ponemos de manifiesto esa diferencia comparando la ecuación en derivadas parciales (9.1) con la ecuación diferencial ordinaria

$$f'(x) = 0. \tag{9.3}$$

La función más general que satisface (9.3) es $f(x) = C$, siendo C una constante arbitraria. Es decir, el espacio-solución de (9.3) es uni-dimensional. Pero la función de tipo más general que satisface (9.1) es

$$f(x, y) = g(y),$$

donde g es una función cualquiera de y. Puesto que g es arbitraria podemos obtener con facilidad un conjunto infinito de soluciones independientes. Por ejemplo, podemos tomar $g(y) = e^{cy}$ y hacer que c varíe en todo el campo real. Así pues, el espacio solución de (9.1) es de *infinitas* dimensiones.

En ciertos aspectos este ejemplo refleja lo que en general ocurre. Algunas veces en la resolución de una ecuación en derivadas parciales de primer orden, se precisa una integración para hacer desaparecer cada derivada parcial. Entonces se introduce un función arbitraria en la solución. Esto ocurre en un espacio de soluciones de infinitas dimensiones.

En muchos problemas que incluyen ecuaciones en derivadas parciales es preciso seleccionar del conjunto de soluciones una solución particular que satisfaga una o más condiciones auxiliares. Como es lógico la naturaleza de esas condiciones tiene un efecto profundo en la existencia o en la unicidad de las soluciones. En este libro no se llegará a un estudio sistemático de tales problemas. En cambio, trataremos algunos casos particulares para ilustrar las ideas introducidas en el capítulo 8.

9.2 Ecuación en derivadas parciales de primer orden con coeficientes constantes

Consideremos la ecuación en derivadas parciales de primer orden

$$3\frac{\partial f(x, y)}{\partial x} + 2\frac{\partial f(x, y)}{\partial y} = 0. \tag{9.4}$$

Todas las soluciones de esta ecuación pueden encontrarse mediante consideraciones geométricas. Expresemos el primer miembro como un producto escalar, y pongamos la ecuación en la forma

$$(3\boldsymbol{i} + 2\boldsymbol{j}) \cdot \nabla f(x, y) = 0.$$

Esto nos dice que el vector gradiente $\nabla f(x, y)$ es ortogonal al vector $3i + 2j$ en cada punto (x, y). Pero sabemos también que $\nabla f(x, y)$ es ortogonal a las curvas de nivel de f. Luego esas curvas de nivel deben ser rectas paralelas a $3i + 2j$. Es decir, las curvas de nivel de f son las rectas

$$2x - 3y = c.$$

Por lo tanto $f(x, y)$ es constante cuando $2x - 3y$ es constante. Esto sugiere que

$$f(x, y) = g(2x - 3y) \tag{9.5}$$

para alguna función g.

Comprobemos ahora que, para cada función diferenciable g, el campo escalar f definido por (9.5) satisface realmente (9.4). Utilizando la regla de la cadena para calcular las derivadas parciales de f encontramos

$$\frac{\partial f}{\partial x} = 2g'(2x - 3y), \qquad \frac{\partial f}{\partial y} = -3g'(2x - 3y),$$

$$3\frac{\partial f}{\partial x} + 2\frac{\partial f}{\partial y} = 6g'(2x - 3y) - 6g'(2x - 3y) = 0.$$

Por consiguiente, f satisface (9.4).

Recíprocamente, podemos demostrar que toda función f diferenciable que satisfaga (9.4) necesariamente debe ser de la forma (9.5) para una cierta g. Para ello, introduzcamos un cambio lineal de variables,

$$x = Au + Bv, \qquad y = Cu + Dv. \tag{9.6}$$

Éste transforma $f(x, y)$ en una función de u y v, sea ésta

$$h(u, v) = f(Au + Bv, Cu + Dv).$$

Elegiremos las constantes A, B, C, D de modo que h satisfaga la ecuación más sencilla

$$\frac{\partial h(u, v)}{\partial u} = 0. \tag{9.7}$$

La resolveremos y demostraremos que f tiene la forma deseada.

Con la regla de la cadena encontramos

$$\frac{\partial h}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} = \frac{\partial f}{\partial x}A + \frac{\partial f}{\partial y}C.$$

Puesto que f satisface (9.4) tenemos $\partial f/\partial y = -(3/2)(\partial f/\partial x)$, por lo que la ecuación toma la forma

$$\frac{\partial h}{\partial u} = \frac{\partial f}{\partial x}\left(A - \frac{3}{2}C\right).$$

Por consiguiente, h satisfará (9.7) si elegimos $A = \frac{3}{2}C$. Tomando $A = 3$ y $C = 2$ encontramos

(9.8) $$x = 3u + Bv, \qquad y = 2u + Dv.$$

Con esta elección de A y C, la función h satisface (9.7), así que $h(u, v)$ es una función sólo de v,

$$h(u, v) = g(v)$$

para una cierta función g. Para expresar v en función de x e y eliminamos u en (9.8) y obtenemos $2x - 3y = (2B - 3D)v$. Elijamos ahora B y D de manera que $2B - 3D = 1$, por ejemplo $B = 2$, $D = 1$. Para estos valores la transformación (9.6) es no singular; tenemos $v = 2x - 3y$, y por tanto

$$f(x, y) = h(u, v) = g(v) = g(2x - 3y).$$

Esto demuestra que toda función diferenciable f solución de (9.4) tiene la forma (9.5).

El mismo tipo de razonamiento demuestra el siguiente teorema para las ecuaciones de primer orden con coeficientes constantes.

TEOREMA 9.1. *Si g es una función diferenciable en $\mathbf{R}^1$ y f es el campo escalar definido en $\mathbf{R}^2$ por medio de la ecuación*

(9.9) $$f(x, y) = g(bx - ay),$$

en la que a y b son constantes, no simultáneamente nulas, entonces f satisface la ecuación en derivadas parciales de primer orden

(9.10) $$a\,\frac{\partial f(x, y)}{\partial x} + b\,\frac{\partial f(x, y)}{\partial y} = 0$$

en todo $\mathbf{R}^2$. Recíprocamente, toda solución diferenciable de (9.10) *tiene necesariamente la forma* (9.9) *para una cierta g.*

9.3 Ejercicios

En este conjunto de ejercicios puede suponerse la diferenciabilidad de todas las funciones que se consideran.

1. Determinar la solución de la ecuación en derivadas parciales

$$4\frac{\partial f(x,y)}{\partial x}+3\frac{\partial f(x,y)}{\partial y}=0$$

que satisfaga la condición $f(x,0)=\operatorname{sen} x$ para todo x.

2. Determinar la solución de la ecuación en derivadas parciales

$$5\frac{\partial f(x,y)}{\partial x}-2\frac{\partial f(x,y)}{\partial y}=0$$

que satisfaga las condiciones $f(0,0)=0$, y $D_1f(x,0)=e^x$ para todo x.

3. a) Si $u(x,y)=f(xy)$, demostrar que u satisface la ecuación en derivadas parciales

$$x\frac{\partial u}{\partial x}-y\frac{\partial u}{\partial y}=0.$$

Hallar una solución tal que $u(x,x)=x^4e^{x^2}$ para todo x.

b) Si $v(x,y)=f(x/y)$ para $y\neq 0$, demostrar que v satisface la ecuación en derivadas parciales

$$x\frac{\partial v}{\partial x}+y\frac{\partial v}{\partial y}=0.$$

Hallar una solución tal que $v(1,1)=2$ y $D_1v(x,1/x)=1/x$ para todo $x\neq 0$.

4. Si $g(u,v)$ satisface la ecuación en derivadas parciales

$$\frac{\partial^2 g(u,v)}{\partial u\,\partial v}=0,$$

demostrar que $g(u,v)=\varphi_1(u)+\varphi_2(v)$, donde $\varphi_1(u)$ es una función sólo de u y $\varphi_2(v)$ lo es únicamente de v.

5. Supóngase que f satisface la ecuación en derivadas parciales

$$\frac{\partial^2 f}{\partial x^2}-2\frac{\partial^2 f}{\partial x\,\partial y}-3\frac{\partial^2 f}{\partial y^2}=0.$$

Introducir el cambio lineal de variables $x=Au+Bv$, $y=Cu+Dv$, siendo A, B, C, D constantes, y póngase $g(u,v)=f(Au+Bv,\ Cu+Dv)$. Calcular valores enteros no nulos de A, B, C, D para los que g satisfaga $\partial^2 g/(\partial u\,\partial v)=0$. Resolver esta ecuación corres-

pondiente a g y determinar luego f. (Supóngase la igualdad de las derivadas parciales mixtas.)

6. Una función u está definida mediante una ecuación de la forma

$$u(x, y) = xyf\left(\frac{x+y}{xy}\right).$$

Demostrar que u satisface la ecuación diferencial de la forma

$$x^2 \frac{\partial u}{\partial x} - y^2 \frac{\partial u}{\partial y} = G(x, y)u,$$

y hallar $G(x, y)$.

7. La sustitución $x = e^s$, $y = e^t$ transforma $f(x, y)$ en $g(s, t)$, siendo $g(s, t) = f(e^s, e^t)$. Si se sabe que f satisface la ecuación en derivadas parciales

$$x^2 \frac{\partial^2 f}{\partial x^2} + y^2 \frac{\partial^2 f}{\partial y^2} + x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = 0,$$

demostrar que g satisface la ecuación en derivadas parciales

$$\frac{\partial^2 g}{\partial s^2} + \frac{\partial^2 g}{\partial t^2} = 0.$$

8. Sea f un campo escalar diferenciable en un conjunto abierto S de $\mathbf{R}^n$. Decimos que f es *homogénea de grado p* en S si

$$f(t\boldsymbol{x}) = t^p f(\boldsymbol{x})$$

para todo $t > 0$ y todo $\boldsymbol{x}$ de S para los que $t\boldsymbol{x} \in S$. Demostrar que para un campo escalar homogéneo de grado p se tiene

$$\boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) = p f(\boldsymbol{x}) \quad \text{para cada } \boldsymbol{x} \text{ de } S.$$

Éste es el llamado *teorema de Euler para las funciones homogéneas.* Si $\boldsymbol{x} = (x_1, \ldots, x_n)$ puede expresarse del siguiente modo

$$x_1 \frac{\partial f}{\partial x_1} + \cdots + x_n \frac{\partial f}{\partial x_n} = p f(x_1, \ldots, x_n).$$

[*Indicación*: Para $\boldsymbol{x}$, fijo, defínase $g(t) = f(t\boldsymbol{x})$ y calcular $g'(1)$.]

9. Demostrar el recíproco del teorema de Euler. Esto es, si f satisface $\boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) = pf(\boldsymbol{x})$ para todo $\boldsymbol{x}$ en un conjunto abierto S, f debe ser homogéneo de grado p en S. [*Indicación*: Para $\boldsymbol{x}$, fijo, defínase $g(t) = f(t\boldsymbol{x}) - t^p f(\boldsymbol{x})$ y calcular $g'(t)$.]

10. Demostrar la siguiente extensión del teorema de Euler para funciones homogéneas de grado p en el caso de dos dimensiones. (Supóngase la igualdad de las derivadas parciales mixtas.)

$$x^2 \frac{\partial^2 f}{\partial x^2} + 2xy \frac{\partial^2 f}{\partial x\, \partial y} + y^2 \frac{\partial^2 f}{\partial y^2} = p(p-1)f.$$

9.4 La ecuación de ondas uni-dimensional

Imaginemos una cuerda tensa de longitud infinita a lo largo del eje x y que puede vibrar en el plano xy. Designemos con $y = f(x, t)$ el desplazamiento vertical de la cuerda en el punto x en el instante t. Supongamos que, en el instante $t = 0$, la cuerda está desplazada tomando la forma de una curva $y = F(x)$. En la figura 9.1 a) se representa un ejemplo. Las figuras 9.1 b) y c) muestran las posibles curvas de desplazamiento para valores posteriores de t. Consideremos el desplazamiento $f(x, t)$ como una función incógnita de x y t que hay que determinar. Un modelo matemático para este problema (sugerido por consideraciones físicas que aquí no comentaremos) es la ecuación en derivadas parciales

$$\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2},$$

en la que c es una constante positiva que depende de las características físicas de la cuerda. Esta ecuación es la llamada *ecuación de ondas uni-dimensional*. La resolveremos teniendo en cuenta ciertas condiciones auxiliares.

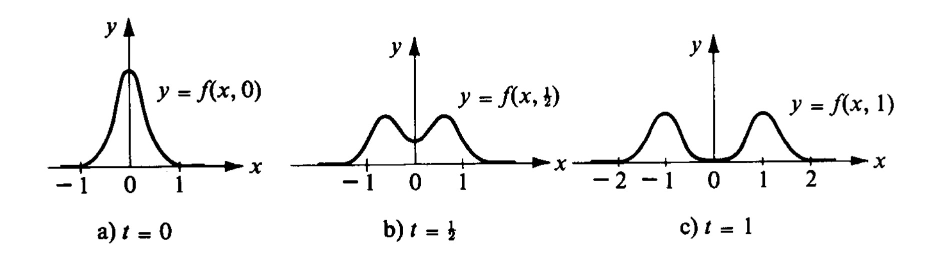


FIGURA 9.1 *Curva de desplazamiento $y = f(x, t)$ para varios valores de t.*

Puesto que el desplazamiento inicial es la curva dada $y = F(x)$, vamos a buscar una solución que satisfaga la condición

$$f(x, 0) = F(x).$$

Supondremos también que $\partial y/\partial t$, la velocidad de desplazamiento vertical, está dada en el instante $t = 0$, a saber

$$D_2 f(x, 0) = G(x),$$

siendo G una función dada. Parece razonable pensar que esta información debería bastar para determinar el subsiguiente movimiento de la cuerda. Demostraremos que esto es cierto, efectivamente, determinando la función f por medio de F y G. La solución se expresa en una forma dada por Jean d'Alembert (1717-1783), matemático y filósofo francés.

TEOREMA 9.2. SOLUCIÓN DE D'ALEMBERT DE LA ECUACIÓN DE ONDAS. *Dadas las funciones F y G tales que G es derivable y F dos veces derivable en* $\mathbf{R}^1$. *La función f dada por la fórmula*

$$f(x, t) = \frac{F(x + ct) + F(x - ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} G(s)\, ds \tag{9.11}$$

satisface la ecuación de ondas uni-dimensional

$$\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2} \tag{9.12}$$

y las condiciones iniciales

$$f(x, 0) = F(x), \qquad D_2 f(x, 0) = G(x). \tag{9.13}$$

Recíprocamente, cualquier función f con derivadas parciales iguales que satisfaga (9.12) *y* (9.13) *tiene necesariamente la forma* (9.11).

Demostración. Comprobar que la función f dada por (9.11) satisface la ecuación de ondas y las condiciones iniciales, es un ejercicio que dejamos al lector. Demostraremos el recíproco.

Un modo de hacerlo consiste en suponer que f es una solución de la ecuación de ondas, introducir un cambio lineal de variables,

$$x = Au + Bv, \qquad t = Cu + Dv,$$

que transforma $f(x, t)$ en una función de u y v,

$$g(u, v) = f(Au + Bv, Cu + Dv),$$

y elegir las constantes A, B. C, D de modo que g satisfaga la ecuación más sencilla

$$\frac{\partial^2 g}{\partial u\, \partial v} = 0.$$

Resolviendo esta ecuación encontramos que $g(u, v) = \varphi_1(u) + \varphi_2(v)$, donde $\varphi_1(u)$ es una función de u solamente y donde $\varphi_2(v)$ sólo lo es de v. Las constantes A, B, C, D pueden elegirse de manera que $u = x + ct$, $v = x - ct$, de donde obtenemos

$$f(x, t) = \varphi_1(x + ct) + \varphi_2(x - ct). \tag{9.14}$$

Hecho esto utilizamos las condiciones iniciales (9.13) para determinar las funciones φ_1 y φ_2 por medio de las funciones dadas F y G.

Obtendremos (9.14) por medio de otro método que utiliza el teorema 9.1 y evita el cambio de variables. Escribimos primero la ecuación de ondas en la forma

$$L_1(L_2 f) = 0, \tag{9.15}$$

siendo L_1 y L_2 operadores diferenciales lineales de primer orden dados por

$$L_1 = \frac{\partial}{\partial t} - c\frac{\partial}{\partial x}, \qquad L_2 = \frac{\partial}{\partial t} + c\frac{\partial}{\partial x}.$$

Sea f una solución de (9.15) y pongamos

$$u(x, t) = L_2 f(x, t).$$

La ecuación (9.15) establece que u satisface la ecuación de primer orden $L_1(u) = 0$. Luego, según el teorema 9.1 tenemos

$$u(x, t) = \varphi(x + ct)$$

para una cierta función φ. Sea Φ una función primitiva de φ, tal como $\Phi(y) = \int_0^y \varphi(s)\,ds$, y pongamos

$$v(x, t) = \frac{1}{2c}\Phi(x + ct).$$

Demostraremos que $L_2(v) = L_2(f)$. Tenemos

$$\frac{\partial v}{\partial x} = \frac{1}{2c}\Phi'(x+ct) \qquad \text{y} \qquad \frac{\partial v}{\partial t} = \frac{1}{2}\Phi'(x+ct),$$

así que

$$L_2 v = \frac{\partial v}{\partial t} + c\,\frac{\partial v}{\partial x} = \Phi'(x+ct) = \varphi(x+ct) = u(x,t) = L_2 f.$$

Es decir, la diferencia $f - v$ satisface la ecuación de primer orden

$$L_2(f-v) = 0.$$

Según el teorema 9.1 debe ser $f(x, t) - v(x, t) = \psi(x - ct)$ para una cierta función ψ. Por consiguiente

$$f(x,t) = v(x,t) + \psi(x-ct) = \frac{1}{2c}\Phi(x+ct) + \psi(x-ct).$$

Esto demuestra (9.14) poniendo $\varphi_1 = \dfrac{1}{2c}\Phi$ y $\varphi_2 = \psi$.

Utilicemos ahora las condiciones iniciales (9.13) para determinar las funciones φ_1 y φ_2 por medio de las funciones dadas F y G. La relación $f(x, 0) = F(x)$ implica

$$\varphi_1(x) + \varphi_2(x) = F(x). \tag{9.16}$$

La otra condición inicial, $D_2 f(x, 0) = G(x)$, implica

$$c\varphi_1'(x) - c\varphi_2'(x) = G(x). \tag{9.17}$$

Derivando (9.16) obtenemos

$$\varphi_1'(x) + \varphi_2'(x) = F'(x). \tag{9.18}$$

Resolviendo (9.17) y (9.18) respecto a $\varphi_1'(x)$ y $\varphi_2'(x)$ encontramos

$$\varphi_1'(x) = \frac{1}{2}F'(x) + \frac{1}{2c}G(x), \qquad \varphi_2'(x) = \frac{1}{2}F'(x) - \frac{1}{2c}G(x).$$

Integrando esas relaciones llegamos a

$$\varphi_1(x) - \varphi_1(0) = \frac{F(x) - F(0)}{2} + \frac{1}{2c}\int_0^x G(s)\,ds,$$

$$\varphi_2(x) - \varphi_2(0) = \frac{F(x) - F(0)}{2} - \frac{1}{2c}\int_0^x G(s)\,ds.$$

Sustituyamos en la primera ecuación x por $x + ct$ y en la segunda x por $x - ct$. Luego sumemos las dos ecuaciones que resulten y teniendo en cuenta que $\varphi_1(0) + \varphi_2(0) = F(0)$ obtenemos

$$f(x, t) = \varphi_1(x + ct) + \varphi_2(x - ct) = \frac{F(x + ct) + F(x - ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} G(s)\,ds.$$

Esto completa la demostración.

EJEMPLO. Supongamos que el desplazamiento inicial viene dado por la fórmula

$$F(x) = \begin{cases} 1 + \cos \pi x & \text{para} \quad -1 \leq x \leq 1, \\ 0 & \text{para} \quad |x| \geq 1. \end{cases}$$

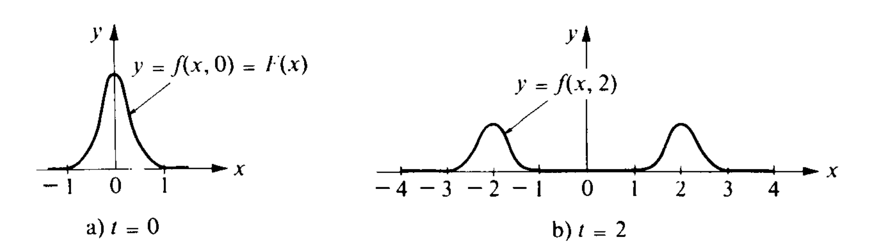


FIGURA 9.2 *Solución de la ecuación de ondas representada para los valores $t = 0$ y $t = 2$.*

La gráfica de F se muestra en las figuras 9.1 a) y 9.2 a). Supongamos que la velocidad inicial $G(x) = 0$ para todo x. Entonces la solución que resulta de la ecuación de ondas viene dada por la fórmula

$$f(x, t) = \frac{F(x + ct) + F(x - ct)}{2}.$$

Las figuras 9.1 y 9.2 representan la curva $y = f(x, t)$ para varios valores de t. Las figuras ponen de manifiesto que la solución de la ecuación de ondas es una combinación de dos ondas estacionarias, una que se desplaza hacia la derecha y la otra a la izquierda, cada una con velocidad c.

En los ejercicios que siguen se dan ejemplos en los que se utiliza la regla de la cadena en el estudio de las ecuaciones en derivadas parciales.

9.5 Ejercicios

Podemos suponer en estos ejercicios la diferenciabilidad de todas las funciones que se consideran.

1. Si k es una constante positiva y $g(x, t) = \frac{1}{2}x/\sqrt{kt}$, ponemos

$$f(x, t) = \int_0^{g(x,t)} e^{-u^2}\, du .$$

a) Demostrar que

$$\frac{\partial f}{\partial x} = e^{-g^2} \frac{\partial g}{\partial x} \quad \text{y} \quad \frac{\partial f}{\partial t} = e^{-g^2} \frac{\partial g}{\partial t} .$$

b) Demostrar que f satisface la ecuación en derivadas parciales

$$k \frac{\partial^2 f}{\partial x^2} = \frac{\partial f}{\partial t} \quad \textit{(ecuación del calor).}$$

2. Consideremos un campo escalar f definido en $\mathbf{R}^2$ tal que $f(x, y)$ dependa sólo de la distancia r del punto (x, y) al origen, $f(x, y) = g(r)$, siendo $r = (x^2 + y^2)^{1/2}$.

a) Demostrar que para $(x, y) \neq (0, 0)$ tenemos

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{1}{r} g'(r) + g''(r).$$

b) Supongamos además que f satisface la *ecuación de Laplace,*

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0,$$

para todo $(x, y) \neq (0, 0)$. Demostrar, aplicando la parte a) que $f(x, y) = a \log (x^2 + y^2) + b$ para $(x, y) \neq (0, 0)$, siendo a y b constantes.

3. Repetir el ejercicio 2 en el caso n-dimensional, siendo $n \geq 3$. Esto es, supóngase que $f(x) = f(x_1, \ldots, x_n) = g(r)$, siendo $r = \|x\|$. Demostrar que

$$\frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2} = \frac{n-1}{r} g'(r) + g''(r)$$

para $x \neq O$. Si f satisface la *ecuación de Laplace n-dimensional,*

$$\frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2} = 0,$$

para todo $x \neq O$, deducir que $f(x) = a\,\|x\|^{2-n} + b$ para $x \neq O$, siendo a y b constantes.

Observación: El operador lineal ∇^2 definido por la ecuación

$$\nabla^2 f = \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}$$

se llama *laplaciana n-dimensional.*

4. *Laplaciana bi-dimensional en coordenadas polares.* La introducción de coordenadas polares $x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$, transforma $f(x, y)$ en $g(r, \theta)$. Comprobar las siguientes fórmulas:

a) $\|\nabla f(r\cos\theta, r\,\mathrm{sen}\,\theta)\|^2 = \left(\frac{\partial g}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial g}{\partial \theta}\right)^2.$

b) $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} + \frac{1}{r}\frac{\partial g}{\partial r}.$

5. *Laplaciana tridimensional en coordenadas esféricas.* La introducción de coordenadas esféricas

$$x = \rho\cos\theta\,\mathrm{sen}\,\varphi, \qquad y = \rho\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\varphi, \qquad z = \rho\cos\varphi,$$

transforma $f(x, y, z)$ en $F(\rho, \theta, \varphi)$. Este ejercicio indica cómo hay que proceder para expresar la laplaciana $\nabla^2 f$ en función de las derivadas parciales de F.

a) Introducir primero las coordenadas polares $x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$ para transformar $f(x, y, z)$ en $g(r, \theta, z)$. Utilizar el ejercicio 4 para demostrar que

$$\nabla^2 f = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{\partial^2 g}{\partial z^2}.$$

b) Transformar luego $g(r, \theta, z)$ en $F(\rho, \theta, \varphi)$ tomando $z = \rho\cos\varphi$, $r = \rho\,\mathrm{sen}\,\varphi$. Obsérvese que, salvo un cambio de notación, esta transformación es la misma que la utilizada en la parte a). Deducir

$$\nabla^2 f = \frac{\partial^2 F}{\partial \rho^2} + \frac{2}{\rho}\frac{\partial F}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 F}{\partial \varphi^2} + \frac{\cos\varphi}{\rho^2\,\mathrm{sen}\,\varphi}\frac{\partial F}{\partial \varphi} + \frac{1}{\rho^2\,\mathrm{sen}\,\varphi}\frac{\partial^2 F}{\partial \theta^2}.$$

6. Este ejercicio pone de manifiesto que la ecuación de Legendre aparece cuando buscamos soluciones de la ecuación de Laplace que tengan una forma especial. Sea f un campo escalar que satisfaga la ecuación de Laplace tri-dimensional, $\nabla^2 f = 0$. Introduzcamos coordenadas esféricas como en el ejercicio 5 y pongamos $F(\rho, \theta, \varphi) = f(x, y, z)$.

a) Supongamos que buscamos soluciones f de la ecuación de Laplace tales que $F(\rho, \theta, \varphi)$ sean independientes de θ y tengan la forma particular $F(\rho, \theta, \varphi) = \rho^n G(\varphi)$. Demostrar que f satisface la ecuación de Laplace si G satisface la ecuación de segundo orden

$$\frac{d^2G}{d\varphi^2} + \cot \varphi \frac{dG}{d\varphi} + n(n+1)G = 0.$$

b) El cambio de variable $x = \cos \varphi$ ($\varphi = \arccos x$, $-1 \leq x \leq 1$) transforma $G(\varphi)$ en $g(x)$. Demostrar que g satisface la ecuación de Legendre

$$(1 - x^2)\frac{d^2g}{dx^2} - 2x\frac{dg}{dx} + n(n+1)g = 0.$$

7. *Ecuación bi-dimensional de ondas.* Una membrana delgada y flexible está extendida sobre el plano xy y puede vibrar. Designemos con $z = f(x, y, t)$ el desplazamiento vertical de la membrana en el punto (x, y) en el instante t. Ciertas consideraciones físicas sugieren que f satisface la ecuación bi-dimensional de ondas

$$\frac{\partial^2 f}{\partial t^2} = c^2\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right),$$

en la que c es una constante positiva que depende de las características físicas de la membrana. Este ejercicio revela una conexión entre esta ecuación y la ecuación diferencial de Bessel.

a) Introducir coordenadas polares $x = r\cos\theta$, $y = r \operatorname{sen} \theta$, y pongamos $F(r, \theta, t) = f(r\cos\theta, r \operatorname{sen}\theta, t)$. Si f satisface la ecuación de ondas demostrar que F satisface la ecuación

$$\frac{\partial^2 F}{\partial t^2} = c^2\left(\frac{\partial^2 F}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 F}{\partial \theta^2} + \frac{1}{r}\frac{\partial F}{\partial r}\right).$$

b) Si $F(r, \theta, t)$ es independiente de θ, $F(r, \theta, t) = \varphi(r, t)$ la ecuación a) se reduce a

$$\frac{\partial^2 \varphi}{\partial t^2} = c^2\left(\frac{\partial^2 \varphi}{\partial r^2} + \frac{1}{r}\frac{\partial \varphi}{\partial r}\right).$$

Sea ahora φ una solución tal que $\varphi(r, t)$ se escinde en el producto de una función de r por una función de t, $\varphi(r, t) = R(r)T(t)$. Demostrar que cada una de las funciones R y T satisface una ecuación diferencial lineal de segundo orden.

c) Si la función T del apartado b) es periódica de período $2\pi/c$, demostrar que R satisface la ecuación de Bessel $r^2R'' + rR' + r^2R = 0$.

9.6 Derivación de funciones definidas implícitamente

Algunas superficies en el espacio de tres dimensiones se representan por ecuaciones cartesianas de la forma

$$F(x, y, z) = 0.$$

Una ecuación como ésta, decimos, proporciona una *representación implícita* de la superficie. Por ejemplo, la ecuación $x^2 + y^2 + z^2 - 1 = 0$ representa la superficie de una esfera unidad con centro en el origen. Algunas veces es posible despejar en la ecuación $F(x, y, z) = 0$ una de las variables en función de las otras dos, por ejemplo z en función de x e y. Esto nos conduce a una o varias ecuaciones de la forma

$$z = f(x, y).$$

Para la esfera tenemos dos soluciones,

$$z = \sqrt{1 - x^2 - y^2} \qquad \text{y} \qquad z = -\sqrt{1 - x^2 - y^2},$$

una representa la semiesfera superior, la otra la semiesfera inferior.

En general no resulta sencillo obtener una fórmula explícita para z en función de x e y. Por ejemplo, no hay un método para despejar con facilidad z en la ecuación $y^2 + xz + z^2 - e^z - 4 = 0$. Sin embargo, utilizando en forma adecuada la regla de la cadena se pueden deducir varias propiedades de las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ sin un conocimiento explícito de $f(x, y)$. En esta sección se expone dicho método.

Supongamos que existe una función $f(x, y)$ tal que

$$F[x, y, f(x, y)] = 0 \tag{9.19}$$

para todo (x, y) en un cierto conjunto abierto S, y que no sea posible disponer de fórmulas explícitas para calcular $f(x, y)$. Expresemos ésta diciendo que $F(x, y, z) = 0$ define z *implícitamente* como función de x e y, y escribamos

$$z = f(x, y).$$

Introduzcamos ahora una función auxiliar g definida en S como sigue:

$$g(x, y) = F[x, y, f(x, y)].$$

La ecuación (9.19) establece que $g(x, y) = 0$ en S; luego las derivadas parciales $\partial g/\partial x$ y $\partial g/\partial y$ son también 0 en S. Pero también podemos calcular esas derivadas parciales mediante la regla de la cadena. Para ello escribamos

$$g(x, y) = F[u_1(x, y), u_2(x, y), u_3(x, y)],$$

siendo $u_1(x, y) = x$, $u_2(x, y) = y$, y $u_3(x, y) = f(x, y)$. La regla de la cadena nos da las fórmulas

$$\frac{\partial g}{\partial x} = D_1F\frac{\partial u_1}{\partial x} + D_2F\frac{\partial u_2}{\partial x} + D_3F\frac{\partial u_3}{\partial x}, \qquad \frac{\partial g}{\partial y} = D_1F\frac{\partial u_1}{\partial y} + D_2F\frac{\partial u_2}{\partial y} + D_3F\frac{\partial u_3}{\partial y},$$

en las que cada derivada parcial D_kF está calculada en $(x, y, f(x, y))$. Puesto que tenemos

$$\frac{\partial u_1}{\partial x} = 1, \qquad \frac{\partial u_2}{\partial x} = 0, \qquad \frac{\partial u_3}{\partial x} = \frac{\partial f}{\partial x}, \qquad \text{y} \qquad \frac{\partial g}{\partial x} = 0,$$

la primera de las ecuaciones anteriores se transforma en

$$D_1F + D_3F\frac{\partial f}{\partial x} = 0.$$

Resolviéndola respecto a $\partial f/\partial x$ obtenemos

(9.20) $$\frac{\partial f}{\partial x} = -\frac{D_1F[x, y, f(x, y)]}{D_3F[x, y, f(x, y)]}$$

en los puntos en los que $D_3F[x, y, f(x, y)] \neq 0$. Mediante un razonamiento parecido obtenemos una fórmula análoga para $\partial f/\partial y$:

(9.21) $$\frac{\partial f}{\partial y} = -\frac{D_2F[x, y, f(x, y)]}{D_3F[x, y, f(x, y)]}$$

en los puntos en los que $D_3F[x, y, f(x, y)] \neq 0$. Ordinariamente esas fórmulas se escriben más brevemente así:

$$\frac{\partial f}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z}, \qquad \frac{\partial f}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z}.$$

EJEMPLO. Supongamos que la ecuación $y^2 + xz + z^z - e^z - c = 0$ define z como función de x e y, sea ésta $z = f(x, y)$. Hallar un valor de la constante c para el cual $f(0, e) = 2$, y calcular las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ en el punto $(x, y) = (0, e)$.

Solución. Cuando $x = 0$, $y = e$ y $z = 2$ la ecuación toma la forma $e^2 + 4 - e^2 - c = 0$, y ésta se satisface para $c = 4$. Sea $F(x, y, z) = y^2 + xz + z^2 - e^z - 4$. De (9.20) y (9.21) obtenemos

$$\frac{\partial f}{\partial x} = -\frac{z}{x + 2z - e^z}, \qquad \frac{\partial f}{\partial y} = -\frac{2y}{x + 2z - e^z}.$$

Cuando $x = 0$, $y = e$, y $z = 2$ encontramos $\partial f/\partial x = 2/(e^2 - 4)$ y $\partial f/\partial y = 2e/(e^2 - 4)$. Obsérvese que hemos podido calcular las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ utilizando tan sólo el valor de $f(x, y)$ en el punto $(0, e)$.

La discusión anterior puede extenderse a funciones de más de dos variables.

TEOREMA 9.3. *Si F es un campo escalar diferenciable en un conjunto abierto T de $\mathbf{R}^n$ y si suponemos que la ecuación*

$$F(x_1, \ldots, x_n) = 0$$

define implícitamente x_n como función diferenciable de $x_1, \ldots, x_{n-1}$

$$x_n = f(x_1, \ldots, x_{n-1}),$$

para todos los puntos $(x_1, \ldots, x_{n-1})$ en un cierto conjunto S de $\mathbf{R}^{n-1}$, entonces para cada $k = 1, 2, \ldots, n - 1$, la derivada parcial $D_k f$ viene dada por la fórmula

(9.22) $$D_k f = -\frac{D_k F}{D_n F}$$

en los puntos en los que $D_n F \neq 0$. Las derivadas parciales $D_k F$ y $D_n F$ que aparecen en (9.22) *están calculadas en el punto $(x_1, x_2, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1}))$.*

La demostración es una extensión directa del razonamiento utilizado para deducir las ecuaciones (9.20) y (9.21) y la dejamos para el lector.

La discusión puede generalizarse en otro sentido. Supongamos dos superficies con las representaciones implícitas siguientes:

(9.23) $$F(x, y, z) = 0, \qquad G(x, y, z) = 0.$$

Si esas superficies se cortan a lo largo de una curva C, es posible obtener una

representación paramétrica de C resolviendo las dos ecuaciones (9.23) simultáneamente respecto a dos de las variables en función de la tercera, por ejemplo x e y en función de z. Supongamos que sea posible despejar x e y y que las soluciones vengan dadas por las ecuaciones

$$x = X(z), \qquad y = Y(z)$$

para todo z en un cierto intervalo abierto (a, b) Entonces cuando x e y se reemplazan por $X(z)$ e $Y(z)$, respectivamente, las dos ecuaciones (9.23) se satisfacen idénticamente. Esto es, podemos escribir $F[X(z), Y(z), z] = 0$ y $G[X(z), Y(z), z] = 0$ para todo z de (a, b). Volviendo a usar la regla de la cadena, podemos calcular las derivadas $X'(z)$ e $Y'(z)$ sin un conocimiento explícito de $X(z)$ e $Y(z)$. Introduzcamos para ello dos nuevas funciones f y g por medio de las ecuaciones

$$f(z) = F[X(z), Y(z), z] \quad \text{y} \quad g(z) = G[X(z), Y(z), z].$$

Entonces $f(z) = g(z) = 0$ para todo z de (a, b) y por tanto las derivadas $f'(z)$ y $g'(z)$ también son cero en (a, b). Con la regla de la cadena esas derivadas se expresan con las fórmulas

$$f'(z) = \frac{\partial F}{\partial x} X'(z) + \frac{\partial F}{\partial y} Y'(z) + \frac{\partial F}{\partial z}, \qquad g'(z) = \frac{\partial G}{\partial x} X'(z) + \frac{\partial G}{\partial y} Y'(z) + \frac{\partial G}{\partial z}.$$

Puesto que $f'(z)$ y $g'(z)$ son ambas cero podemos determinar $X'(z)$ e $Y'(z)$ resolviendo el siguiente sistema de ecuaciones *lineales:*

$$\frac{\partial F}{\partial x} X'(z) + \frac{\partial F}{\partial y} Y'(z) = -\frac{\partial F}{\partial z},$$

$$\frac{\partial G}{\partial x} X'(z) + \frac{\partial G}{\partial y} Y'(z) = -\frac{\partial G}{\partial z}.$$

En aquellos puntos en los que el determinante del sistema no es nulo, el sistema tiene una sola solución que, mediante la regla de Cramer, se puede expresar así:

$$(9.24) \qquad X'(z) = -\frac{\begin{vmatrix} \frac{\partial F}{\partial z} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial z} & \frac{\partial G}{\partial y} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{vmatrix}}, \qquad Y'(z) = -\frac{\begin{vmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial z} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial z} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{vmatrix}}.$$

Los determinantes que aparecen en (9.24) son determinantes de matrices de Jacobi y se llaman *determinantes jacobianos.* Con frecuencia se emplea una notación especial para los determinantes jacobianos. Escribimos

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} = \det \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}.$$

Con esta notación, las fórmulas (9.24) pueden expresarse más brevemente en la forma

$$(9.25) \qquad X'(z) = \frac{\partial(F, G)/\partial(y, z)}{\partial(F, G)/\partial(x, y)}, \qquad Y'(z) = \frac{\partial(F, G)/\partial(z, x)}{\partial(F, G)/\partial(x, y)}.$$

(El signo menos se ha incorporado a los numeradores permutando las columnas.)

El método se puede extender a casos más generales en los que se dan m ecuaciones con n variables, siendo $n > m$ obteniéndose m variables en función de las $n - m$ restantes. Las derivadas parciales de las nuevas funciones así definidas se pueden expresar como cocientes de determinantes de Jacobi, generalizando así (9.25). En el ejercicio 3 de la sección 9.8 se da un ejemplo en el que $m = 2$ y $n = 4$.

9.7 Ejemplos resueltos

En esta sección ilustramos algunos de los conceptos de la anterior resolviendo algunos problemas relativos a funciones definidas implícitamente.

EJEMPLO 1. Supongamos que la ecuación $g(x, y) = 0$ determina y como función derivable de x, sea ésta $y = Y(x)$ para todo x en un cierto intervalo (a, b). Expresar la derivada $Y'(x)$ en función de las derivadas parciales de g.

Solución. Sea $G(x) = g[x, Y(x)]$ para x en (a, b). Entonces la ecuación $g(x, y) = 0$ implica $G(x) = 0$ en (a, b). En virtud de la regla de la cadena tenemos

$$G'(x) = \frac{\partial g}{\partial x} \cdot 1 + \frac{\partial g}{\partial y} Y'(x),$$

de la que obtenemos

$$Y'(x) = -\frac{\partial g/\partial x}{\partial g/\partial y} \tag{9.26}$$

en los puntos x de (a, b) en los que $\partial g/\partial y \neq 0$. Las derivadas parciales $\partial g/\partial x$ y $\partial g/\partial y$ vienen dadas por las fórmulas $\partial g/\partial x = D_1 g[x, Y(x)]$ y $\partial g/\partial y = D_2 g[x, Y(x)]$.

EJEMPLO 2. Cuando se elimina y entre las dos ecuaciones $z = f(x, y)$ y $g(x, y) = 0$, el resultado puede expresarse en la forma $z = h(x)$. Expresar la derivada $h'(x)$ en función de las derivadas parciales de f y g.

Solución. Supongamos que la ecuación $g(x, y) = 0$ puede resolverse respecto a y en función de x y que una solución sea $y = Y(x)$ para todos los valores de x de un cierto intervalo abierto (a, b). Entonces la función h será

$$h(x) = f[x, Y(x)] \quad \text{si} \quad x \in (a, b).$$

Aplicando la regla de la cadena tenemos

$$h'(x) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} Y'(x).$$

Con la ecuación (9.26) del ejemplo 1 obtenemos la fórmula

$$h'(x) = \frac{\dfrac{\partial g}{\partial y}\dfrac{\partial f}{\partial x} - \dfrac{\partial f}{\partial y}\dfrac{\partial g}{\partial x}}{\dfrac{\partial g}{\partial y}}.$$

Las derivadas parciales del segundo miembro están calculadas en el punto $(x, Y(x))$. Obsérvese que el numerador también puede expresarse como un determinante jacobiano, resultando

$$h'(x) = \frac{\partial(f, g)/\partial(x, y)}{\partial g/\partial y}.$$

EJEMPLO 3. Las dos ecuaciones $2x = v^2 - u^2$ e $y = uv$ definen u y v como funciones de x e y. Hallar las fórmulas correspondientes a $\partial u/\partial x, \partial u/\partial y, \partial v/\partial x, \partial v/\partial y$.

Solución. Si mantenemos fija y y derivamos las dos ecuaciones citadas con

respecto a x, recordando que u y v son funciones de x e y, obtenemos

$$2 = 2v\frac{\partial v}{\partial x} - 2u\frac{\partial u}{\partial x} \qquad \text{y} \qquad 0 = u\frac{\partial v}{\partial x} + v\frac{\partial u}{\partial x}.$$

Resolviendo estas ecuaciones respecto a $\partial u/\partial x$ y $\partial v/\partial x$ encontramos

$$\frac{\partial u}{\partial x} = -\frac{u}{u^2 + v^2} \qquad \text{y} \qquad \frac{\partial v}{\partial x} = \frac{v}{u^2 + v^2}.$$

Por otra parte, si mantenemos fija x y derivamos las dos ecuaciones dadas respecto a y obtenemos las ecuaciones

$$0 = 2v\frac{\partial v}{\partial y} - 2u\frac{\partial u}{\partial y} \qquad \text{y} \qquad 1 = u\frac{\partial v}{\partial y} + v\frac{\partial u}{\partial y}.$$

De este sistema de dos ecuaciones obtenemos

$$\frac{\partial u}{\partial y} = \frac{v}{u^2 + v^2} \qquad \text{y} \qquad \frac{\partial v}{\partial y} = \frac{u}{u^2 + v^2}.$$

EJEMPLO 4. Sea u una función de x e y definida por la ecuación

$$u = F(x + u, yu).$$

Hallar $\partial u/\partial x$ y $\partial u/\partial y$ en función de las derivadas parciales de F.

Solución. Supongamos que $u = g(x, y)$ para todo (x, y) en un cierto conjunto abierto S. Sustituyendo $g(x,y)$ por u en la ecuación original obtenemos

$$g(x, y) = F[u_1(x, y), u_2(x, y)], \tag{9.27}$$

en donde $u_1(x, y) = x + g(x, y)$ y $u_2(x, y) = y\ g(x, y)$. Mantengamos ahora y fija y derivemos ambos miembros de (9.27) respecto a x, empleando la regla de la cadena en el segundo miembro, con lo que obtenemos

$$\frac{\partial g}{\partial x} = D_1F\frac{\partial u_1}{\partial x} + D_2F\frac{\partial u_2}{\partial x}. \tag{9.28}$$

Pero $\partial u_1/\partial x = 1 + \partial g/\partial x$, y $\partial u_2/\partial x = y\,\partial g/\partial x$. Luego (9.28) se convierte en

$$\frac{\partial g}{\partial x} = D_1F\cdot\left(1 + \frac{\partial g}{\partial x}\right) + D_2F\cdot\left(y\frac{\partial g}{\partial x}\right).$$

Resolviendo esta ecuación respecto a $\partial g/\partial x$ (y poniendo $\partial u/\partial x$ en lugar de $\partial g/\partial x$) obtenemos

$$\frac{\partial u}{\partial x} = \frac{-D_1F}{D_1F + y\,D_2F - 1}.$$

Del mismo modo encontramos

$$\frac{\partial g}{\partial y} = D_1F\,\frac{\partial u_1}{\partial y} + D_2F\,\frac{\partial u_2}{\partial y} = D_1F\,\frac{\partial g}{\partial y} + D_2F\left(y\,\frac{\partial g}{\partial y} + g(x, y)\right).$$

Esto nos conduce a la ecuación

$$\frac{\partial u}{\partial y} = \frac{-g(x, y)\,D_2F}{D_1F + y\,D_2F - 1}.$$

Las derivadas parciales D_1F y D_2F están calculadas en el punto $(x + g(x, y), yg(x, y))$.

EJEMPLO 5. Cuando u se elimina entre las dos ecuaciones $x = u + v$ e $y = uv^2$ llegamos a una ecuación de la forma $F(x, y, v) = 0$ que define implícitamente v como función de x e y, sea $v = h(x, y)$. Demostrar que

$$\frac{\partial h}{\partial x} = \frac{h(x, y)}{3h(x, y) - 2x}$$

y encontrar una fórmula análoga para $\partial h/\partial y$.

Solución. Eliminando u entre las dos ecuaciones dadas, obtenemos la relación

$$xv^2 - v^3 - y = 0.$$

Sea F la función definida por la ecuación

$$F(x, y, v) = xv^2 - v^3 - y.$$

Podemos aplicar ahora lo dicho en la sección 9.6 y escribir

(9.29) $$\frac{\partial h}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial v} \quad \text{y} \quad \frac{\partial h}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial v}.$$

Pero $\partial F/\partial x = v^2$, $\partial F/\partial v = 2xv - 3v^2$ y $\partial F/\partial y = -1$. Luego las fórmulas (9.29) se convierten en

$$\frac{\partial h}{\partial x} = -\frac{v^2}{2xv - 3v^2} = -\frac{v}{2x - 3v} = \frac{h(x, y)}{3h(x, y) - 2x}$$

$$\frac{\partial h}{\partial y} = -\frac{-1}{2xv - 3v^2} = \frac{1}{2xh(x, y) - 3h^2(x, y)}.$$

EJEMPLO 6. La ecuación $F(x, y, z) = 0$ define implícitamente z como función de x e y, sea $z = f(x, y)$. Suponiendo que $\partial^2F/(\partial x\, \partial z) = \partial^2F/(\partial z\, \partial x)$, demostrar que tenemos

$$(9.30) \qquad \frac{\partial^2 f}{\partial x^2} = -\frac{\left(\dfrac{\partial^2 F}{\partial z^2}\right)\left(\dfrac{\partial F}{\partial x}\right)^2 - 2\left(\dfrac{\partial^2 F}{\partial x\, \partial z}\right)\left(\dfrac{\partial F}{\partial z}\right)\left(\dfrac{\partial F}{\partial x}\right) + \left(\dfrac{\partial F}{\partial z}\right)^2\left(\dfrac{\partial^2 F}{\partial x^2}\right)}{\left(\dfrac{\partial F}{\partial z}\right)^3},$$

donde las derivadas parciales del segundo miembro están calculadas en $(x, y, f(x, y))$.

Solución. Según la fórmula (9.20) de la sección 9.6 tenemos,

$$(9.31) \qquad \frac{\partial f}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z}.$$

Hay que recordar que este cociente significa en realidad

$$-\frac{D_1F[x, y, f(x, y)]}{D_3F[x, y, f(x, y)]}.$$

Introduzcamos $G(x, y) = D_1F[x, y, f(x, y)]$ y $H(x, y) = D_3F[x, y, f(x, y)]$. Nos proponemos calcular la derivada parcial respecto a x del cociente

$$\frac{\partial f}{\partial x} = -\frac{G(x, y)}{H(x, y)},$$

manteniendo la y fija. Aplicando la regla de la derivación de un cociente resulta,

$$(9.32) \qquad \frac{\partial^2 f}{\partial x^2} = -\frac{H\dfrac{\partial G}{\partial x} - G\dfrac{\partial H}{\partial x}}{H^2}.$$

Puesto que G y H son funciones compuestas, usamos la regla de la cadena para calcular las derivadas parciales $\partial G/\partial x$ y $\partial H/\partial x$. Para $\partial G/\partial x$ tenemos,

$$\begin{aligned}\frac{\partial G}{\partial x} &= D_1(D_1F)\cdot 1 + D_2(D_1F)\cdot 0 + D_3(D_1F)\cdot\frac{\partial f}{\partial x}\\ &= \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial z\,\partial x}\frac{\partial f}{\partial x}.\end{aligned}$$

Análogamente, encontramos

$$\begin{aligned}\frac{\partial H}{\partial x} &= D_1(D_3F)\cdot 1 + D_2(D_3F)\cdot 0 + D_3(D_3F)\cdot\frac{\partial f}{\partial x}\\ &= \frac{\partial^2 F}{\partial x\,\partial z} + \frac{\partial^2 F}{\partial z^2}\frac{\partial f}{\partial x}.\end{aligned}$$

Sustituyendo en (9.32) y reemplazando $\partial f/\partial x$ por el cociente (9.31) obtenemos la fórmula (9.30).

9.8 Ejercicios

En los ejercicios de esta sección se supone la existencia y la continuidad de todas las derivadas que intervienen.

1. Las dos ecuaciones $x + y = uv$ y $xy = u - v$ definen x e y como funciones implícitas de u y v, sean éstas $x = X(u, v)$ e $y = Y(u, v)$. Demostrar que $\partial X/\partial u = (xv - 1)/(x - y)$ si $x \neq y$, y hallar fórmulas parecidas para $\partial X/\partial v$, $\partial Y/\partial u$, $\partial Y/\partial v$.
2. Las dos ecuaciones $x + y = uv$ y $xy = u - v$ definen x y v como funciones de u e y, sean éstas $x = X(u, y)$ y $v = V(u, y)$. Demostrar que $\partial X/\partial u = (u + v)/(1 + yu)$ si $1 + yu \neq 0$, y hallar las fórmulas de $\partial X/\partial y$, $\partial V/\partial u$, $\partial V/\partial y$.
3. Las dos ecuaciones $F(x, y, u, v) = 0$ y, $G(x, y, u, v) = 0$ determinan x e y como funciones implícitas de u y v, sean éstas $x = X(u, v)$ e $y = Y(u, v)$. Demostrar que

$$\frac{\partial X}{\partial u} = \frac{\partial(F, G)/\partial(y, u)}{\partial(F, G)/\partial(x, y)}$$

en los puntos en los que el jacobiano $\partial(F, G)/\partial(x, y) \neq 0$, y hallar las fórmulas para las derivadas parciales $\partial X/\partial v$, $\partial Y/\partial u$ y $\partial Y/\partial v$.

4. La intersección de las dos superficies dadas por las ecuaciones cartesianas $2x^2 + 3y^2 - z^2 = 25$ y $x^2 + y^2 = z^2$ contiene una curva C que pasa por el punto $P = (\sqrt{7}, 3, 4)$. Esas ecuaciones pueden resolverse respecto a x e y en función de z con lo que se obtiene la representación paramétrica de C con z como parámetro.
a) Hallar un vector unitario T tangente a C en el punto P sin utilizar el conocimiento explícito de la representación paramétrica.
b) Confrontar el resultado del apartado anterior a) mediante la representación paramétrica de C con z como parámetro.

5. Las tres ecuaciones $F(u, v) = 0$, $u = xy$ y $v = \sqrt{x^2 + z^2}$ definen una superficie en el espacio xyz. Hallar un vector normal a esa superficie en el punto $x = 1$, $y = 1$, $z = \sqrt{3}$ si se sabe que $D_1F(1, 2) = 1$ y $D_2F(1, 2) = 2$.
6. Las tres ecuaciones

$$x^2 - y\cos(uv) + z^2 = 0,$$

$$x^2 + y^2 - \operatorname{sen}(uv) + 2z^2 = 2,$$

$$xy - \operatorname{sen} u \cos v + z = 0,$$

definen x, y, z como funciones de u y v. Calcular las *derivadas parciales* $\partial x/\partial u$ y $\partial x/\partial v$ en el punto $x = y = 1$, $u = \pi/2$, $v = 0$, $z = 0$.
7. La ecuación $f(y/x, z/x) = 0$ define z implícitamente como función de x e y, sea esa función $z = g(x, y)$. Demostrar que

$$x\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial y} = g(x, y)$$

en los puntos en los que $D_2f[y/x, g(x, y)/x]$ es distinta de cero.
8. Sea F una función real de dos variables reales y supongamos que las derivadas parciales D_1F y D_2F son siempre distintas de cero. Sea u otra función real de dos variables reales tales que las derivadas parciales $\partial u/\partial x$ y $\partial u/\partial y$ están ligadas por la ecuación $F(\partial u/\partial x, \partial u/\partial y) = 0$. Demostrar que existe una constante n tal que

$$\frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial y^2} = \left(\frac{\partial^2 u}{\partial x\, \partial y}\right)^n,$$

y encontrar n. Suponer que $\partial^2 u/(\partial x\, \partial y) = \partial^2 u/(\partial y\, \partial x)$.
9. La ecuación $x + z + (y + z)^2 = 6$ define z como función implícita de x e y, sea $z = f(x, y)$. Calcular las derivadas parciales $\partial f/\partial x$, $\partial f/\partial y$, y $\partial^2 f/(\partial x\, \partial y)$ en función de x, y, y z.
10. La ecuación $\operatorname{sen}(x + y) + \operatorname{sen}(y + z) = 1$ define z como función implícita de x e y, sea $z = f(x, y)$. Calcular la derivada segunda $D_{1,2}f$ en función de x, y, y z.
11. La ecuación $F(x + y + z, x^2 + y^2 + z^2) = 0$ define z como función implícita de x e y, sea $z = f(x, y)$. Determinar las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ en función de las parciales D_1F y D_2F.
12. Sean f y g dos funciones de una variable real y definamos $F(x, y) = f[x + g(y)]$. Hallar las fórmulas correspondientes a todas las derivadas parciales de F de primero y segundo orden, expresadas en función de las derivadas de f y g. Comprobar la relación

$$\frac{\partial F}{\partial x}\frac{\partial^2 F}{\partial x\, \partial y} = \frac{\partial F}{\partial y}\frac{\partial^2 F}{\partial x^2}.$$

9.9 Máximos, mínimos y puntos de ensilladura

Una superficie definida explícitamente por una ecuación de la forma

$z = f(x, y)$ puede considerarse como una superficie de nivel del campo escalar F definido por la ecuación

$$F(x, y, z) = f(x, y) - z .$$

Si f es diferenciable, el gradiente de ese campo viene dado por el vector

$$\nabla F = \frac{\partial f}{\partial x} \boldsymbol{i} + \frac{\partial f}{\partial y} \boldsymbol{j} - \boldsymbol{k} .$$

La ecuación lineal que representa el plano tangente en un punto $P_1 = (x_1, y_1, z_1)$ puede escribirse en la forma

$$z - z_1 = A(x - x_1) + B(y - y_1),$$

en la que

$$A = D_1 f(x_1, y_1) \qquad \text{y} \qquad B = D_2 f(x_1, y_1).$$

Cuando los dos coeficientes A y B son nulos, el punto P_1 se llama *punto estacionario* de la superficie y el punto (x_1, y_1) se llama *punto estacionario* o *crítico* de la función f. El plano tangente en un punto estacionario es horizontal. Generalmente los puntos estacionarios de una superficie se clasifican en tres categorías: máximos, mínimos y puntos de ensilladura. Si la superficie se imagina como un terreno montañoso, esas categorías corresponden, respectivamente, a las cumbres, a los fondos de los valles y a los puertos.

Los conceptos de máximos, mínimos y puntos de ensilladura, se pueden introducir para campos escalares cualesquiera definidos en subconjuntos de $\mathbf{R}^n$.

DEFINICIÓN. *Se dice que un campo escalar f tiene un máximo absoluto en un punto $\boldsymbol{a}$ de un conjunto S de $\mathbf{R}^n$ si*

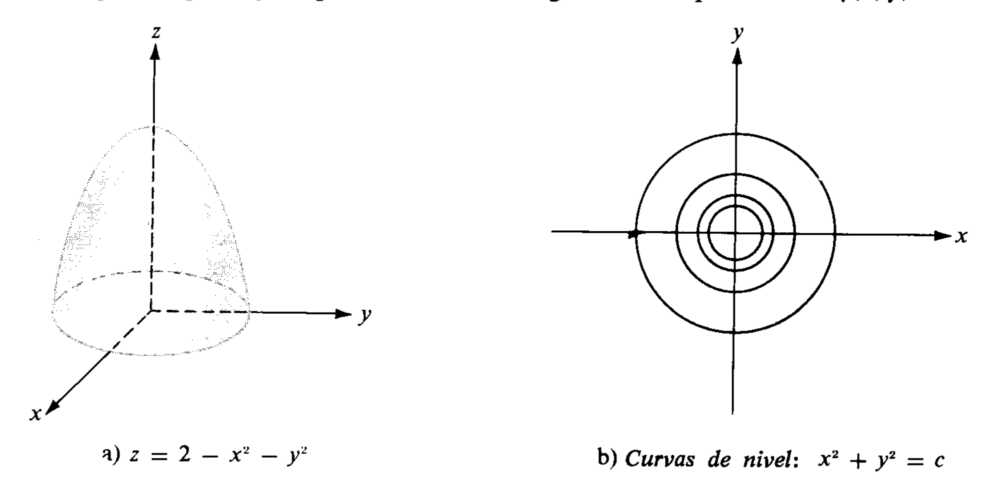
$$f(\boldsymbol{x}) \leq f(\boldsymbol{a}) \tag{9.33}$$

para todo $\boldsymbol{x}$ de S. El número $f(\boldsymbol{a})$ se llama máximo absoluto de f en S. Se dice que la función f tiene un máximo relativo en $\boldsymbol{a}$ si la desigualdad (9.33) *se satisface para todo $\boldsymbol{x}$ de una cierta n-bola $B(\boldsymbol{a})$ contenida en S.*

Dicho de otro modo, un máximo relativo en $\boldsymbol{a}$ es el máximo absoluto en un cierto entorno de $\boldsymbol{a}$. El *mínimo absoluto* y el *mínimo relativo* se definen de modo parecido, empleando la desigualdad opuesta a la (9.33). Algunas veces se emplean los adjetivos *global* y *local* en lugar de *absoluto* y *relativo* respectivamente.

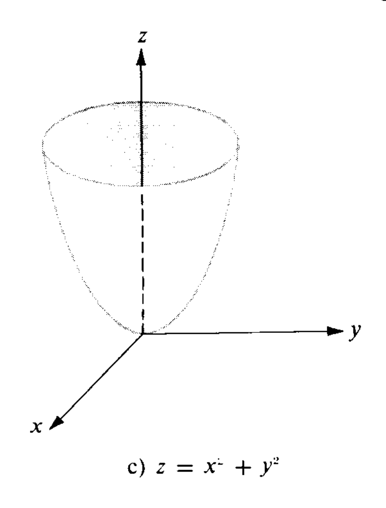
DEFINICIÓN. *Un número que sea máximo relativo o mínimo relativo de f se llama extremo de f.*

Si f tiene un extremo en un punto interior $\boldsymbol{a}$ y es diferenciable en él, todas las derivadas parciales de primer orden $D_1 f(\boldsymbol{a}), \ldots, D_n f(\boldsymbol{a})$ deben ser cero. Es decir, $\nabla f(\boldsymbol{a}) = \boldsymbol{O}$ (Esto se puede probar fácilmente manteniendo fijo cada componente y reduciendo el problema al caso uni-dimensional). En el caso $n = 2$, esto significa que hay un plano horizontal tangente a la superficie $z = f(x, y)$ en el



a) $z = 2 - x^2 - y^2$ b) *Curvas de nivel:* $x^2 + y^2 = c$

Ejemplo 1. Máximo relativo en el origen.



c) $z = x^2 + y^2$

Ejemplo 2. Mínimo relativo en el origen.

FIGURA 9.3 *Ejemplos 1 y 2.*

punto $(\boldsymbol{a}, f(\boldsymbol{a}))$. Por otra parte, es sencillo encontrar ejemplos en los que la anulación de todas las derivadas parciales en $\boldsymbol{a}$ no implica necesariamente un extremo en $\boldsymbol{a}$. Esto sucede en los llamados *puntos de ensilladura* que se definen del modo siguiente.

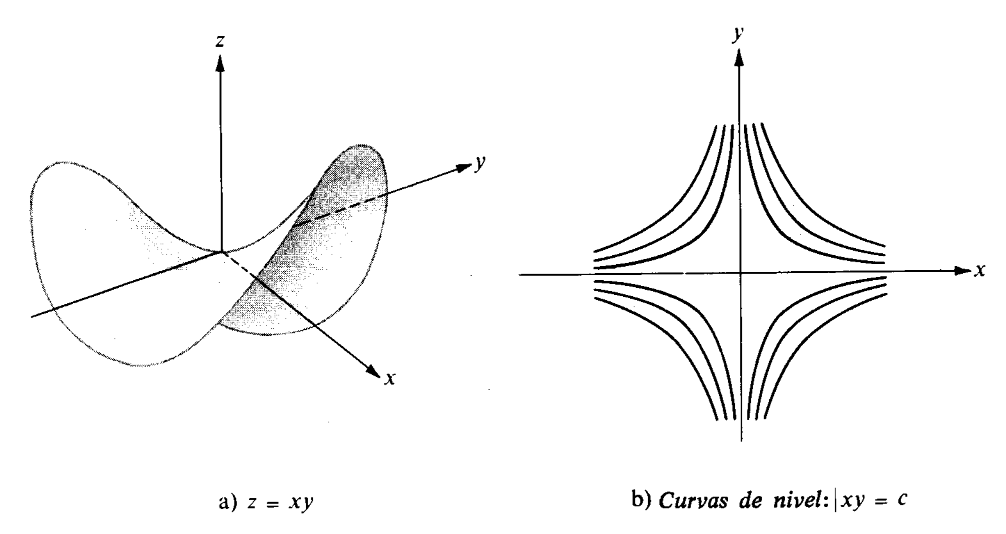
DEFINICIÓN. *Supongamos que f sea diferenciable en $\boldsymbol{a}$. Si $\nabla f(\boldsymbol{a}) = \boldsymbol{O}$ el punto $\boldsymbol{a}$ se llama punto estacionario de f. Un punto estacionario se llama de ensilladura si toda n-bola $B(\boldsymbol{a})$ contiene puntos $\boldsymbol{x}$ tales que $f(\boldsymbol{x}) < f(\boldsymbol{a})$ y otros para los que $f(\boldsymbol{x}) > f(\boldsymbol{a})$.*

La definición es análoga a la del caso uni-dimensional en el que los puntos estacionarios de una función se clasifican en máximos, mínimos y puntos de inflexión. En los ejemplos que siguen se consideran varios tipos de puntos estacionarios. En cada caso el punto estacionario que se considera es el origen.

EJEMPLO 1. *Máximo relativo.* $z = f(x, y) = 2 - x^2 - y^2$. Esta superficie es un paraboloide de revolución. En las proximidades del origen tiene la forma indicada en la figura 9.3 a). Sus curvas de nivel son círculos, alguno de los cuales está dibujado en la figura 9.3 b). Puesto que $f(x, y) = 2 - (x^2 + y^2) \leqslant 2 = f(0, 0)$ para todo (x, y), resulta que f no tan sólo tiene en $(0, 0)$ un máximo relativo, sino también un máximo *absoluto* en todo conjunto que contenga el origen. Las dos derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ se anulan en el origen.

EJEMPLO 2. *Mínimo relativo.* $z = f(x, y) = x^2 + y^2$. Este ejemplo, otro paraboloide de revolución, es en esencia el mismo que el ejemplo anterior, salvo que en el origen hay un mínimo en lugar de un máximo. El aspecto de la superficie en las cercanías del origen se aprecia en la figura 9.3 c) y algunas curvas de nivel están dibujadas en la figura 9.3 b).

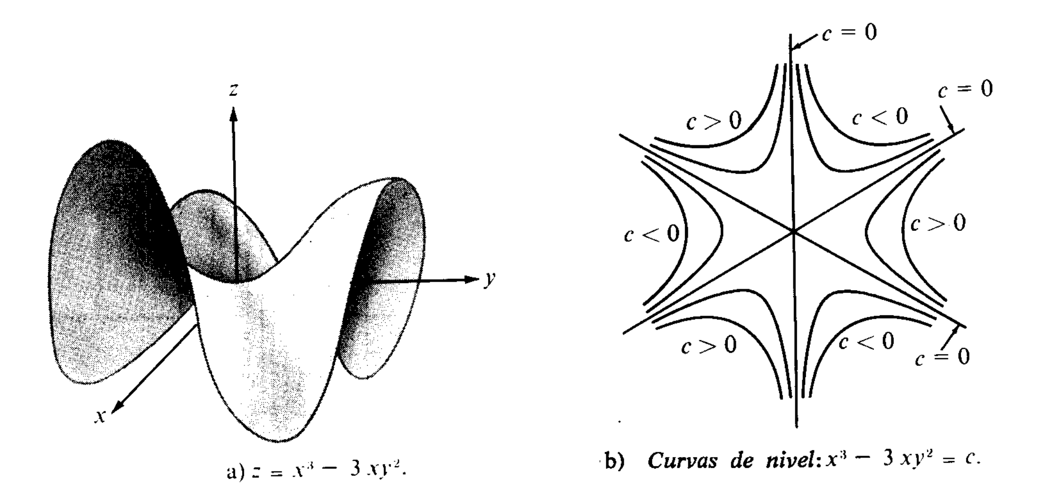
EJEMPLO 3. *Punto de ensilladura.* $z = f(x, y) = xy$. Esta superficie es un paraboloide hiperbólico. Cerca del origen es parecida a una silla de montar, como se ve en la figura 9.4 (a). Las dos derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ son nulas en el origen pero no existe en él ni máximo ni mínimo. En efecto, para puntos (x, y) del primero o tercer cuadrantes, x e y tienen el mismo signo, dándonos $f(x, y) > 0 = f(0, 0)$, mientras que para puntos del segundo y cuarto cuadrantes x e y tienen signos opuestos, y es $f(x, y) < 0 = f(0,0)$. Por consiguiente, en todo entorno del origen hay puntos en los que la función es menor que $f(0, 0)$ y puntos en los que es mayor que $f(0, 0)$, de modo que el origen es un punto de ensilladura. En la figura 9.4 b), se representa también el punto de ensilladura y las curvas de nivel en las proximidades de $(0, 0)$. Esas son hipérbolas que tienen los ejes x e y como asíntotas.



a) $z = xy$ b) *Curvas de nivel*: $xy = c$

FIGURA 9.4 *Ejemplo 3. Punto de ensilladura en el origen.*

EJEMPLO 4. *Punto de ensilladura.* $z = f(x, y) = x^3 - 3xy^2$. Cerca del origen, esta superficie tiene el aspecto de un puerto de montaña entre tres picos. Está representada en la figura 9.5 a). Algunas curvas de nivel se ven en la figura 9.5 b). El punto de ensilladura está en el origen.



a) $z = x^3 - 3xy^2$. b) *Curvas de nivel*: $x^3 - 3xy^2 = c$.

FIGURA 9.5 *Ejemplo 4. Punto de ensilladura en el origen.*

EJEMPLO 5. *Mínimo relativo.* $z = f(x, y) = x^2y^2$. Esta superficie se parece a un valle circundado por cuatro montañas, como sugiere la figura 9.6 a). Existe un mínimo absoluto en el origen, ya que $f(x, y) \geqslant (f(0, 0)$ para todo (x, y). Las curvas de nivel [representadas en la figura 9.6 b)] son hipérbolas cuyas asíntotas son los ejes x e y. Obsérvese que esas curvas de nivel son parecidas a las del ejercicio 3. En este caso, no obstante, la función toma únicamente valores no negativos a lo largo de todas sus curvas de nivel.

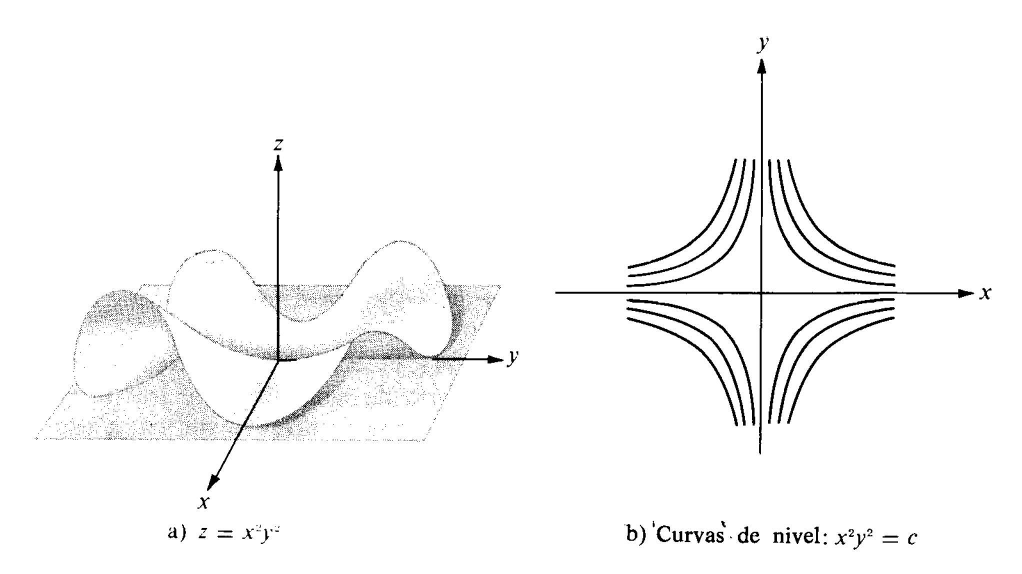


FIGURA 9.6 *Ejemplo 5. Mínimo relativo en el origen.*

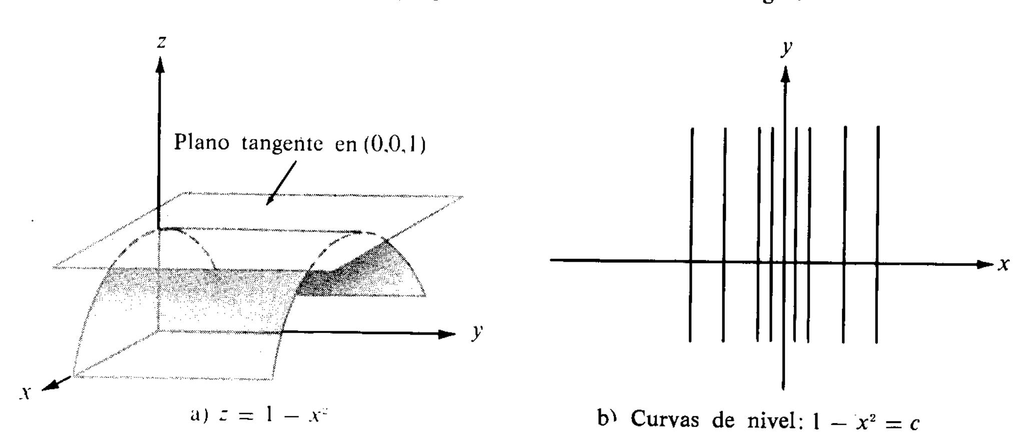


FIGURA 9.7 *Ejemplo 6. Máximo relativo en el origen.*

EJEMPLO 6. *Máximo relativo.* $z = (f(x,y) = 1 - x^2$. En este caso la superficie es un cilindro con generatrices paralelas al eje y, como muestra la figura 9.7 a). Las secciones por planos paralelos al eje x son parábolas. Es evidente que existe un máximo absoluto en el origen debido a que $f(x,y)=1-x^2 \leqslant 1 = f(0,0)$ para todo (x, y). Las curvas de nivel forman una familia de rectas paralelas como se ve en la figura 9.7 b).

9.10 Fórmula de Taylor de segundo orden para campos escalares

Si un campo escalar diferenciable f tiene un punto estacionario en $\boldsymbol{a}$, la naturaleza de éste queda determinada por el signo algebraico de la diferencia $f(\boldsymbol{x}) - f(\boldsymbol{a})$ para $\boldsymbol{x}$ próximo a $\boldsymbol{a}$. Si $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{y}$, tenemos la fórmula de Taylor de primer orden

$$f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{y} + \|\boldsymbol{y}\| \, E(\boldsymbol{a}, \boldsymbol{y}), \text{ donde } E(\boldsymbol{a}, \boldsymbol{y}) \to 0 \text{ cuando } \boldsymbol{y} \to \boldsymbol{O}$$

En un punto estacionario, $\nabla f(\boldsymbol{a}) = \boldsymbol{O}$ y la fórmula de Taylor toma la forma

$$f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a}) = \|\boldsymbol{y}\| \, E(\boldsymbol{a}, \boldsymbol{y}).$$

Para determinar el signo algebraico de $f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a})$ necesitamos más información relativa al término de corrección $\|\boldsymbol{y}\| \, E(\boldsymbol{a}, \boldsymbol{y})$. El teorema que sigue nos dice que si f tiene en $\boldsymbol{a}$, derivadas parciales de segundo orden continuas, el término de corrección o complementario es igual a la forma cuadrática,

$$\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} f(\boldsymbol{a}) y_i y_j$$

más un término de orden menor que $\|\boldsymbol{y}\|^2$. Los coeficientes de la forma cuadrática son las derivadas parciales de segundo orden $D_{ij}f = D_i(D_j f)$, calculadas en $\boldsymbol{a}$. La matriz $n \times n$ de las derivadas segundas $D_{ij}f(\boldsymbol{x})$ es la llamada *matriz hessiana* (*) y se designa por $H(\boldsymbol{x})$. Así pues, tenemos

$$H(\boldsymbol{x}) = [D_{ij} f(\boldsymbol{x})]_{i,j=1}^{n}$$

con tal que existan las derivadas. La forma cuadrática puede escribirse más sencillamente en forma matricial como sigue:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} f(\boldsymbol{a}) y_i y_j = \boldsymbol{y} H(\boldsymbol{a}) \boldsymbol{y}^t,$$

(*) De Ludwig Otto Hesse (1811-1874), matemático alemán autor de muchas contribuciones a la teoría de superficies.

en donde $y = (y_1, \ldots, y_n)$ se considera como una matriz fila $1 \times n$, e y^t es su transpuesta, una matriz columna $n \times 1$. Cuando las derivadas parciales $D_{ij}f$ son continuas tenemos $D_{ij}f = D_{ji}f$ y la matriz $H(a)$ es simétrica.

La fórmula de Taylor que da una aproximación cuadrática para $f(a + y) - f(a)$, toma ahora la siguiente forma.

TEOREMA 9.4. FÓRMULA DE TAYLOR DE SEGUNDO ORDEN PARA CAMPOS ESCALARES. *Si f es un campo escalar con derivadas parciales segundas $D_{ij}f$ continuas en una n-bola $B(a)$, entonces para todo y de $\mathbf{R}^n$ tal que $a + y \in B(a)$ tenemos*

$$(9.34)\quad f(a + y) - f(a) = \nabla f(a) \cdot y + \frac{1}{2!} yH(a + cy)y^t, \qquad \text{donde} \quad 0 < c < 1 .$$

Esto puede escribirse también en la forma

$$(9.35)\qquad f(a + y) - f(a) = \nabla f(a) \cdot y + \frac{1}{2!} yH(a)y^t + \|y\|^2 E_2(a, y),$$

en donde $E_2\ (a, y) \to 0$ cuando $y \to O$.

Demostración. Mantengamos y fijo y definamos $g(u)$ para u real mediante la ecuación

$$g(u) = f(a + uy) \quad \text{para} \quad -1 \leq u \leq 1 .$$

Entonces $f(a + y) - f(a) = g(1) - g(0)$. Demostraremos el teorema aplicando la fórmula de Taylor de segundo orden a g en el intervalo $[0, 1]$. Obtenemos

$$(9.36)\qquad g(1) - g(0) = g'(0) + \frac{1}{2!} g''(c), \qquad \text{donde} \quad 0 < c < 1 .$$

Aquí hemos utilizado para el resto la forma de Lagrange (véase la sección 7.7 del Volumen I).

Puesto que g es una función compuesta dada por $g(u) = f[r(u)]$, siendo $r(u) = a + uy$ podemos calcular su derivada mediante la regla de la cadena. Tenemos $r'(u) = y$ así que la regla de la cadena nos da

$$g'(u) = \nabla f[r(u)] \cdot r'(u) = \nabla f[r(u)] \cdot y = \sum_{j=1}^{n} D_j f[r(u)] y_j ,$$

con tal que $\boldsymbol{r}(u) \in B(\boldsymbol{a})$. En particular, $g'(0) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{y}$. Aplicando una vez más la regla de la cadena encontramos

$$g''(u) = \sum_{i=1}^{n} D_i\left(\sum_{j=1}^{n} D_j f[\boldsymbol{r}(u)] y_j\right) y_i = \sum_{i=1}^{n}\sum_{j=1}^{n} D_{ij} f[\boldsymbol{r}(u)] y_i y_j = \boldsymbol{y} H[\boldsymbol{r}(u)] \boldsymbol{y}^t .$$

Luego $g''(c) = \boldsymbol{y} H(\boldsymbol{a} + c\boldsymbol{y})\boldsymbol{y}^t$, con lo que la ecuación (9.36) se transforma en la (9.34).

Para demostrar (9.35) definamos $E_2(\boldsymbol{a}, \boldsymbol{y})$ por la ecuación

(9.37) $$\|\boldsymbol{y}\|^2 E_2(\boldsymbol{a}, \boldsymbol{y}) = \frac{1}{2!} \boldsymbol{y}\{H(\boldsymbol{a} + c\boldsymbol{y}) - H(\boldsymbol{a})\}\boldsymbol{y}^t \quad \text{si} \quad \boldsymbol{y} \neq \boldsymbol{O},$$

y sea $E_2(\boldsymbol{a}, \boldsymbol{O}) = 0$. Entonces la ecuación (9.34) toma la forma

$$f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{y} + \frac{1}{2!} \boldsymbol{y} H(\boldsymbol{a}) \boldsymbol{y}^t + \|\boldsymbol{y}\|^2 E_2(\boldsymbol{a}, \boldsymbol{y}).$$

Para completar la demostración necesitamos probar que $E_2(\boldsymbol{a}, \boldsymbol{y}) \to 0$ cuando $\boldsymbol{y} \to \boldsymbol{O}$.

De (9.37), obtenemos

$$\|\boldsymbol{y}\|^2 |E_2(\boldsymbol{a}, \boldsymbol{y})| = \frac{1}{2}\left|\sum_{i=1}^{n}\sum_{j=1}^{n} \{D_{ij} f(\boldsymbol{a} + c\boldsymbol{y}) - D_{ij} f(\boldsymbol{a})\} y_i y_j\right|$$

$$\leq \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} |D_{ij} f(\boldsymbol{a} + c\boldsymbol{y}) - D_{ij} f(\boldsymbol{a})| \, \|\boldsymbol{y}\|^2 .$$

Dividiendo por $\|\boldsymbol{y}\|^2$ obtenemos la desigualdad

$$|E_2(\boldsymbol{a}, \boldsymbol{y})| \leq \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} |D_{ij} f(\boldsymbol{a} + c\boldsymbol{y}) - D_{ij} f(\boldsymbol{a})|$$

para $\boldsymbol{y} \neq \boldsymbol{O}$. Puesto que cada derivada parcial segunda $D_{ij}f$ es continua en $\boldsymbol{a}$, tenemos $D_{ij}f(\boldsymbol{a} + c\boldsymbol{y}) \to D_{ij}f(\boldsymbol{a})$ cuando $\boldsymbol{y} \to \boldsymbol{O}$, así que $E_2(\boldsymbol{a}, \boldsymbol{y}) \to 0$ cuando $\boldsymbol{y} \to \boldsymbol{O}$. Esto completa la demostración.

9.11 Determinación de la naturaleza de un punto estacionario por medio de los autovalores de la matriz hessiana

En un punto estacionario tenemos $\nabla f(\boldsymbol{a}) = \boldsymbol{O}$, así que la fórmula de Taylor de la ecuación (9.35) toma la forma

$$f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a}) = \tfrac{1}{2}\boldsymbol{y}H(\boldsymbol{a})\boldsymbol{y}^t + \|\boldsymbol{y}\|^2 E_2(\boldsymbol{a}, \boldsymbol{y}).$$

Puesto que el término de corrección $\|\boldsymbol{y}\|^2 E_2(\boldsymbol{a}, \boldsymbol{y})$ tiende hacia cero más rápidamente que $\|\boldsymbol{y}\|^2$, parece razonable pensar que para $\boldsymbol{y}$ pequeño el signo algebraico de $f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a})$ es el mismo que el de la forma cuadrática$\boldsymbol{y}H(\boldsymbol{a})\boldsymbol{y}^t$;por lo que la naturaleza del punto estacionario podrá determinarse mediante el signo algebraico de la forma cuadrática. Esta sección se dedica a demostrar este hecho.

Damos primero una conexión entre el signo algebraico de una forma cuadrática y sus autovalores.

TEOREMA 9.5. *Sea $A = [a_{ij}]$ una matriz $n\times n$ simétrica, y pongamos*

$$Q(\boldsymbol{y}) = \boldsymbol{y}A\boldsymbol{y}^t = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}y_iy_j .$$

Tenemos entonces:

a) $Q(\boldsymbol{y}) > 0$ *para todo* $\boldsymbol{y} \neq \boldsymbol{O}$ *si y sólo si todos los autovalores de* A *son positivos.*

b) $Q(\boldsymbol{y}) < 0$ *para todo* $\boldsymbol{y} \neq \boldsymbol{O}$ *si y sólo si todos los autovalores de* A *son negativos.*

Observación. En el caso a), la forma cuadrática se llama *definida positiva;* en el caso b) se llama *definida negativa.*

Demostración. En virtud del teorema 5.11 existe una matriz ortogonal C que reduce la forma cuadrática $\boldsymbol{y}A\boldsymbol{y}^t$ a forma diagonal. Esto es

$$Q(\boldsymbol{y}) = \boldsymbol{y}A\boldsymbol{y}^t = \sum_{i=1}^{n} \lambda_i x_i^2 \tag{9.38}$$

donde $\boldsymbol{x} = (x_1, \ldots, x_n)$ es la matriz fila $\boldsymbol{x} = \boldsymbol{y}C$, y $\lambda_1, \ldots, \lambda_n$ son los autovalores de A. Los autovalores son reales puesto que A es simétrica.

Si todos los autovalores son positivos, la ecuación (9.38) pone de manifiesto que $Q(\boldsymbol{y}) > 0$ siempre que $\boldsymbol{x} \neq \boldsymbol{O}$. Pero como $\boldsymbol{x} = \boldsymbol{y}C$, tenemos $\boldsymbol{y} = \boldsymbol{x}C^{-1}$, por lo que $\boldsymbol{x} \neq \boldsymbol{O}$, si y sólo si $\boldsymbol{y} \neq \boldsymbol{O}$. En consecuencia $Q(\boldsymbol{y}) > 0$ para todo $\boldsymbol{y} \neq \boldsymbol{O}$.

Recíprocamente, si $Q(y) > 0$ para todo $y \neq O$ podemos elegir y de modo que $x = yC$ es el k-ésimo vector coordenado e_k. Para este y, la ecuación (9.38) nos da $Q(y) = \lambda_k$, de modo que cada $\lambda_k > 0$. Esto demuestra la parte a). La demostración de b) es análoga.

El teorema que sigue relaciona la naturaleza de un punto estacionario con el signo algebraico de la forma cuadrática $yH(a)y^t$.

TEOREMA 9.6. *Sea f un campo escalar con derivadas parciales segundas continuas $D_{ij}f$ en una n-bola $B(a)$, y designemos con $H(a)$ la matriz hessiana en un punto estacionario a. Tenemos entonces:*

a) *Si todos los autovalores de $H(a)$ son positivos, f tiene un mínimo relativo en a.*

b) *Si todos los autovalores de $H(a)$ son negativos, f tiene un máximo relativo en a.*

c) *Si $H(a)$ tiene autovalores positivos y negativos, f tiene un punto de ensilladura en a.*

Demostración. Pongamos $Q(y) = yH(a)y^t$. La fórmula de Taylor nos da

$$f(a + y) - f(a) = \tfrac{1}{2}Q(y) + \|y\|^2 E_2(a, y), \tag{9.39}$$

en donde $E_2(a, y) \to 0$ cuando $y \to O$. Vamos a demostrar que existe un número positivo r tal que, si $0 < \|y\| < r$, el signo algebraico de $f(a + y) - f(a)$ es el mismo que el de $Q(y)$.

Supongamos primero que todos los autovalores $\lambda_1, \ldots, \lambda_n$ de $H(a)$ son positivos. Sea h el autovalor más pequeño. Si $u < h$, los n números

$$\lambda_1 - u, \ldots, \lambda_n - u$$

son también positivos. Esos números son los autovalores de la matriz real simétrica $H(a) - uI$, siendo I la matriz identidad $n \times n$. Según el teorema 9.5, la forma cuadrática $y[H(a) - uI]y^t$ es definida positiva, y por tanto $y[H(a) - uI]y^t > 0$ para todo $y \neq O$. Por lo tanto

$$yH(a)y^t > y(uI)y^t = u\,\|y\|^2$$

para todo valor real $u < h$. Tomando $u = \frac{1}{2}h$ obtenemos la desigualdad

$$Q(y) > \tfrac{1}{2}h\,\|y\|^2$$

para todo $y \neq O$. Puesto que $E_2(a, y) \to O$ cuando $y \to O$, existe un número

positivo r tal que $|E_2(\boldsymbol{a}, \boldsymbol{y})| < \frac{1}{4}h$ con tal que $0 < \|\boldsymbol{y}\| < r$. Para tal $\boldsymbol{y}$ tenemos

$$0 \leq \|\boldsymbol{y}\|^2 \, |E_2(\boldsymbol{a}, \boldsymbol{y})| < \tfrac{1}{4}h \, \|\boldsymbol{y}\|^2 < \tfrac{1}{2}Q(\boldsymbol{y}),$$

y la fórmula de Taylor (9.39) demuestra que

$$f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a}) \geq \tfrac{1}{2}Q(\boldsymbol{y}) - \|\boldsymbol{y}\|^2 \, |E_2(\boldsymbol{a}, \boldsymbol{y})| > 0.$$

Por consiguiente f tiene un mínimo relativo en $\boldsymbol{a}$, lo que demuestra la parte a). Para probar b) podemos utilizar un razonamiento parecido, o aplicando simplemente la parte a) a $-f$.

Para demostrar c), sean λ_1 y λ_2 dos autovalores de $H(\boldsymbol{a})$ de signos opuestos. Pongamos $h = \min\{|\lambda_1|, |\lambda_2|\}$. Entonces para cada valor real u que satisfaga $-h < u < h$ los números

$$\lambda_1 - u \qquad \text{y} \qquad \lambda_2 - u$$

son autovalores de signos opuestos de la matriz $H(\boldsymbol{a}) - uI$. Por consiguiente, si $u \in (-h, h)$, la forma cuadrática $\boldsymbol{y}[H(\boldsymbol{a}) - uI]\boldsymbol{y}^t$ toma valores positivos y negativos en todo entorno de $\boldsymbol{y} = \boldsymbol{O}$. Elijamos, como antes, $r > 0$ de modo que $|E_2(\boldsymbol{a}, \boldsymbol{y})| < \frac{1}{4}h$ siempre que $0 < \|\boldsymbol{y}\| < r$. Razonando, entonces, como antes vemos que para tal $\boldsymbol{y}$ el signo de $f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a})$ es el mismo que el de $Q(\boldsymbol{y})$. Puesto que para $\boldsymbol{y} \to \boldsymbol{O}$, se presentan valores positivos y negativos, f tiene en $\boldsymbol{a}$ un punto de ensilladura. Esto completa la demostración.

Observación: Si todos los autovalores de $H(\boldsymbol{a})$ son cero, el teorema 9.6 no nos da información relativa al punto estacionario. Se pueden dar criterios, para tratar tales ejemplos, en los que intervienen derivadas de orden superior, pero no los expondremos aquí.

9.12 Criterio de las derivadas segundas para determinar extremos de funciones de dos variables

En el caso $n = 2$ la naturaleza del punto estacionario se puede determinar también mediante el signo algebraico de la derivada segunda $D_{1,1}f(\boldsymbol{a})$ y del determinante de la matriz hessiana.

TEOREMA 9.7. *Sea* $\boldsymbol{a}$ *un punto estacionario de un campo escalar* $f(x_1, x_2)$

con derivadas parciales segundas continuas en una 2-bola $B(\boldsymbol{a})$. Designemos con

$$A = D_{1,1}f(\boldsymbol{a}), \qquad B = D_{1,2}f(\boldsymbol{a}), \qquad C = D_{2,2}f(\boldsymbol{a}),$$

y sea

$$\Delta = \det H(\boldsymbol{a}) = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2.$$

Tenemos entonces:

a) *Si $\Delta < 0$, f tiene un punto de ensilladura en $\boldsymbol{a}$.*
b) *Si $\Delta > 0$ y $A > 0$, f tiene un mínimo relativo en $\boldsymbol{a}$.*
c) *Si $\Delta > 0$ y $A < 0$, f tiene un máximo relativo en $\boldsymbol{a}$.*
d) *Si $\Delta = 0$, el criterio no decide nada.*

Demostración. En este caso la ecuación característica $\det [\lambda I - H(\boldsymbol{a})] = 0$ es cuadrática,

$$\lambda^2 - (A + C)\lambda + \Delta = 0.$$

Los autovalores λ_1, λ_2 están ligados a los coeficientes por las ecuaciones

$$\lambda_1 + \lambda_2 = A + C, \qquad \lambda_1\lambda_2 = \Delta.$$

Si $\Delta < 0$ los autovalores tienen signos opuestos, así que f tiene un punto de ensilladura en $\boldsymbol{a}$, lo que prueba a). Si $\Delta > 0$, los autovalores tienen el mismo signo. En este caso $AC > B^2 \geqslant 0$, así que A y C tienen el mismo signo. Este signo debe ser el de λ_1 y λ_2 ya que $A + C = \lambda_1 + \lambda_2$. Esto demuestra b) y c).

Para demostrar d) nos referiremos a los ejemplos 4 y 5 de la sección 9.9. En ambos tenemos $\Delta = 0$ en el origen. En el ejemplo 4 el origen es un punto de ensilladura, y en el ejemplo 5 es un mínimo relativo.

Aun cuando sea aplicable el teorema 9.7 puede ocurrir que no sea éste el camino más sencillo para determinar la naturaleza de un punto estacionario. Por ejemplo, cuando $f(x, y) = e^{1/g(x,y)}$, en donde $g(x, y) = x^2 + 2 + \cos^2 y - 2 \cos y$, el criterio es aplicable, pero los cálculos son muy largos. En este caso podemos expresar $g(x, y)$ como una suma de cuadrados escribiendo $g(x, y) = 1 + x^2 + (1 - \cos y)^2$. En seguida vemos que f tiene máximos relativos en los puntos en los que $x^2 = 0$ y $(1 - \cos y)^2 = 0$. Estos son los puntos $(0, 2n\pi)$, siendo n un entero cualquiera.

9.13 Ejercicios

En los ejercicios del 1 al 15, identificar y clasificar (si existen) los puntos estacionarios

de las superficies que tienen las ecuaciones cartesianas que se dan.

1. $z = x^2 + (y-1)^2$.
2. $z = x^2 - (y-1)^2$.
3. $z = 1 + x^2 - y^2$.
4. $z = (x - y + 1)^2$.
5. $z = 2x^2 - xy - 3y^2 - 3x + 7y$.
6. $z = x^2 - xy + y^2 - 2x + y$.
7. $z = x^3 - 3xy^2 + y^3$.
8. $z = x^2y^3(6 - x - y)$.
9. $z = x^3 + y^3 - 3xy$.
10. $z = \operatorname{sen} x \cosh y$.
11. $z = e^{2x+3y}(8x^2 - 6xy + 3y^2)$.
12. $z = (5x + 7y - 25)e^{-(x^2+xy+y^2)}$.
13. $z = \operatorname{sen} x \operatorname{sen} y \operatorname{sen}(x + y), \quad 0 \le x \le \pi, 0 \le y \le \pi$.
14. $z = x - 2y + \log\sqrt{x^2 + y^2} + 3\arctan\dfrac{y}{x}, \quad x > 0$.
15. $z = (x^2 + y^2)e^{-(x^2+y^2)}$.
16. Sea $f(x, y) = 3x^4 - 4x^2y + y^2$. Demostrar que sobre toda recta de la forma $y = mx$ la función tiene un mínimo en $(0, 0)$, pero que no existe mínimo relativo en ningún entorno bidimensional del origen. Hacer un dibujo indicando el conjunto de puntos (x, y) en los que $f(x, y) > 0$ y el conjunto en el que $f(x, y) < 0$.
17. Sea $f(x, y) = (3 - x)(3 - y)(x + y - 3)$.
    a) Trazar una figura indicando el conjunto de puntos (x, y) en los que $f(x, y) \ge 0$.
    b) Hallar todos los puntos (x, y) del plano en los que $D_1f(x, y) = D_2f(x, y) = 0$. [*Indicación.* $D_1f(x, y)$ contiene $(3 - y)$ como factor.]
    c) ¿Cuáles de los puntos estacionarios son máximos relativos? ¿Cuáles son mínimos relativos? ¿Cuáles ni una cosa ni otra? Razonar las contestaciones.
    d) ¿Tiene f un mínimo absoluto o un máximo absoluto en todo el plano? Razonar la contestación.
18. Determinar todos los valores extremos absolutos y relativos y los puntos de ensilladura para la función $f(x, y) = xy(1 - x^2 - y^2)$ en el cuadrado $0 \le x \le 1, 0 \le y \le 1$.
19. Determinar las constantes a y b para que la integral

$$\int_0^1 \{ax + b - f(x)\}^2\, dx$$

tome el valor menor posible si a) $f(x) = x^2$; b) $f(x) = (x^2 + 1)^{-1}$.

20. Sea $f(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F$ en donde $A > 0$ y $B^2 < AC$.
    a) Demostrar que existe un punto (x_1, y_1) en el que f tiene un mínimo. [*Indicación.* Transformar la parte cuadrática en una suma de cuadrados.]
    b) Demostrar que $f(x_1, y_1) = Dx_1 + Ey_1 + F$ en ese mínimo.
    c) Demostrar que

$$f(x_1, y_1) = \frac{1}{AC - B^2}\begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix}.$$

21. *Método de los mínimos cuadrados.* Dados n números distintos $x, \ldots, x_n$ y otros n números $y_1, \ldots, y_n$ (no necesariamente distintos), es en general imposible encontrar una recta $f(x) = ax + b$ que pase por todos los puntos (x_i, y_i), esto es, tal que $f(x_i) = y_i$ para cada i. No obstante, podemos encontrar una función lineal con la que el «error cuadrático total»

$$E(a, b) = \sum_{i=1}^{n} [f(x_i) - y_i]^2$$

sea mínimo. Determinar los valores de a y b para que eso ocurra.

22. Extender el método de los mínimos cuadrados a E_3. Esto es, hallar una función lineal $f(x, y) = ax + by + c$ que minimice el error cuadrático total

$$E(a, b, c) = \sum_{i=1}^{n} [f(x_i, y_i) - z_i]^2,$$

donde (x_i, y_i) son n puntos distintos dados y $z_1, \ldots, z_n$ son n números reales dados.

23. Sean $z_1, \ldots, z_n$ n puntos distintos en un m-espacio. Si $\boldsymbol{x} \in \mathbf{R}^m$, definamos

$$f(\boldsymbol{x}) = \sum_{k=1}^{n} \|\boldsymbol{x} - \boldsymbol{z}_k\|^2.$$

Demostrar que f tiene un mínimo en el punto $\boldsymbol{a} = \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{z}_k$ (centroide).

24. Sea $\boldsymbol{a}$ un punto estacionario de un campo escalar f con derivadas parciales segundas en una n-bola $B(\boldsymbol{a})$. Demostrar que f tiene un punto de ensilladura en $\boldsymbol{a}$ si por lo menos dos elementos de la diagonal principal de la matriz hessiana $H(\boldsymbol{a})$ tienen signos opuestos.

25. Comprobar que el campo escalar $f(x, y, z) = x^4 + y^4 + z^4 - 4xyz$ tiene un punto estacionario en $(1, 1, 1)$, y determinar la naturaleza de ese punto estacionario calculando los autovalores de su matriz hessiana.

9.14 Extremos condicionados. Multiplicadores de Lagrange

Iniciamos esta sección con dos ejemplos de problemas de extremos condicionados.

EJEMPLO 1. Dada una superficie S que no pase por el origen, determinar los puntos de S más próximos al origen.

EJEMPLO 2. Si $f(x, y, z)$ representa la temperatura en (x, y, z), determinar los valores máximo y mínimo de la temperatura en una curva dada C del espacio de tres dimensiones.

Ambos ejemplos son casos particulares del siguiente problema general: *Determinar los valores extremos de un campo escalar $f(\boldsymbol{x})$ cuando $\boldsymbol{x}$ tiene la restricción de pertenecer a un subconjunto dado del dominio de f.*

En el ejemplo 1 el campo escalar cuyo mínimo se desea es la función distancia,

$$f(x, y, z) = (x^2 + y^2 + z^2)^{1/2};$$

el subconjunto restringido es la superficie dada S. En el ejemplo 2 tal subconjunto es la curva dada C.

Con frecuencia los problemas de extremos condicionados son muy difíciles; no se conoce un método general para resolverlos con toda generalidad. Se utilizan métodos particulares cuando el subconjunto restringido tiene una estructura sencilla, por ejemplo, si es una superficie como en el ejemplo 1, o una curva como en el ejemplo 2. Esta sección está dedicada al método de los multiplicadores de Lagrange para resolver tales problemas. Ante todo exponemos el método en su forma general, y luego damos argumentos geométricos para comentar su aplicación a los dos ejemplos antes mencionados.

Método de los multiplicadores de Lagrange. Si un campo escalar $f(x_1, \ldots, x_n)$ tiene un extremo relativo cuando está sometido a m condiciones, por ejemplo

$$(9.40) \qquad g_1(x_1, \ldots, x_n) = 0, \qquad \ldots, \qquad g_m(x_1, \ldots, x_n) = 0,$$

siendo $m < n$, existen entonces m escalares $\lambda_1, \ldots, \lambda_m$ tales que

$$(9.41) \qquad \nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_m \nabla g_m$$

en cada punto extremo.

Para determinar los puntos extremos en la práctica consideramos el sistema de $n + m$ ecuaciones formado con las m ecuaciones de condición (9.40) y las n ecuaciones escalares determinadas por la relación vectorial (9.41). Se resuelve el sistema (si ello es posible) respecto a las $n + m$ incógnitas $x_1, \ldots, x_n$ y $\lambda_1, \ldots, \lambda_m$. Los puntos $(x_1, \ldots, x_n)$ en los que se presentan los extremos relativos se encuentran entre las soluciones de aquél sistema.

Los escalares $\lambda_1, \ldots, \lambda_m$ que se introdujeron para ayudarnos a resolver este tipo de problema se denominan *multiplicadores de Lagrange.* Se introduce un multiplicador por cada condición. El campo escalar f y las funciones de condición $g_1, \ldots, g_m$ se suponen diferenciables. El método es válido si el número de condiciones, m, es menor que el número de variables, n, y si no todos los determinantes jacobianos de las funciones de condición con respecto a m de las variables $x_1, \ldots, x_n$ son nulos para los valores extremos que se consideran. La demostración de la validez del método es un resultado importante del cálculo superior y no la expondremos aquí. (Véase el capítulo 7 de la obra del autor *Análisis matemático,* Editorial Reverté, S. A., Barcelona, Buenos Aires, Caracas, México. En lugar de ello daremos unos argumentos geométricos para hacer ver el significado del método y cómo se aplica en los dos ejemplos que al principio se han citado.

Solución geométrica del ejemplo 1. Queremos determinar los puntos de una superficie dada S que están más próximos al origen. Un punto (x, y, z) del espacio

de tres dimensiones está a distancia r del origen si y sólo si está en la esfera

$$x^2 + y^2 + z^2 = r^2.$$

Esta esfera es una superficie de nivel de la función $f(x, y, z) = (x^2 + y^2 + z^2)^{1/2}$ que hay que minimar. Si empezamos con $r = 0$ y aumentamos r hasta que la correspondiente superficie de nivel sea tangente a la superficie dada S, cada punto de contacto será un punto de S lo más próximo al origen.

Para determinar los puntos de contacto suponemos que S está definida por la ecuación cartesiana $g(x, y, z) = 0$. Si S tiene plano tangente en un punto de contacto, dicho plano también debe ser tangente a la superficie de nivel tangente a S en el mismo punto. Por lo tanto el vector gradiente de la superficie $g(x, y, z) = 0$ debe ser paralelo al vector gradiente de la superficie de nivel de contacto $f(x, y, z) = r$. Por tanto existe una constante λ tal que

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$$

en cada punto de contacto. Esa es la ecuación vectorial (9.41) lograda con el método de Lagrange cuando hay una sola condición.

Solución geométrica del ejemplo 2. Queremos obtener los valores extremos

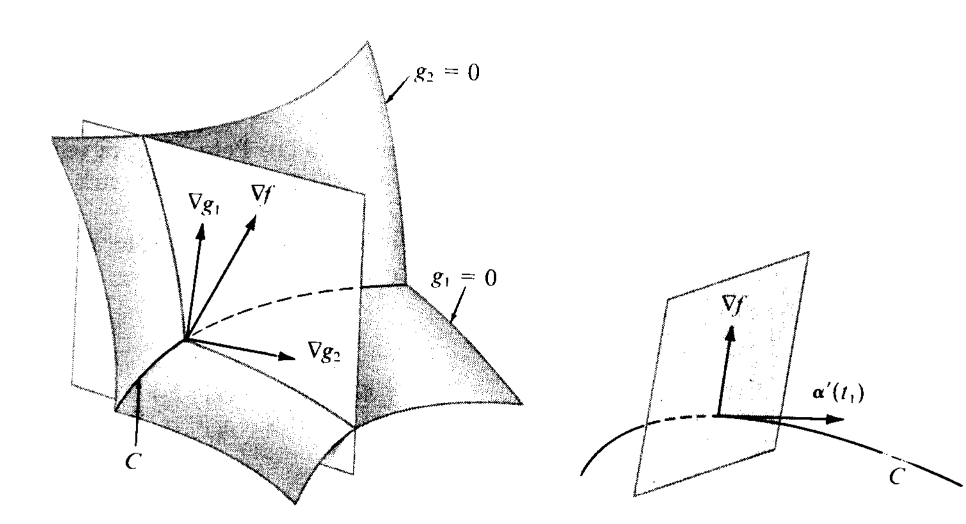


FIGURA 9.8 *Los vectores $\nabla g_1, \nabla g_2$ y ∇f están situados en el mismo plano.*

FIGURA 9.9 *El vector gradiente ∇f está en un plano normal a C.*

de una función que da la temperatura $f(x, y, z)$ sobre una curva dada C. Si consideramos la curva C como la intersección de dos superficies,

$$g_1(x, y, z) = 0 \qquad \text{y} \qquad g_2(x, y, z) = 0,$$

tenemos un problema de extremos con dos condiciones. Los dos vectores gradientes ∇g_1 y ∇g_2 son normales a esas superficies, luego también lo son a la curva C de intersección. (Ver figura 9.8). Seguidamente demostramos que el vector gradiente ∇f de la función temperatura también es normal a C en cada extremo relativo sobre C. Ello implica que ∇f está en el mismo plano que ∇g_1 y ∇g_2; luego si ∇g_1 y ∇g_2 son independientes podemos expresar ∇f como combinación lineal de ∇g_1 y ∇g_2, es decir

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2.$$

Ésta es la ecuación vectorial (9.41) obtenida con el método de Lagrange cuando existen dos condiciones.

Para demostrar que ∇f es normal a C en un punto extremo imaginemos que C esté definida por una función vectorial $\boldsymbol{\alpha}(t)$, variando t en un intervalo $[a, b]$. Sobre la curva C la temperatura se convierte en una función de t, es decir $\varphi(t) = f[\boldsymbol{\alpha}(t)]$. Si φ tiene un extremo relativo en un punto interior t_1 de $[a, b]$ tiene que verificarse $\varphi'(t_1) = 0$. Por otra parte, la regla de la cadena nos dice que $\varphi'(t)$ viene dada por el producto escalar

$$\varphi'(t) = \nabla f[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t).$$

Este producto es nulo en t_1, luego ∇f es perpendicular a $\boldsymbol{\alpha}'(t)$. Pero $\boldsymbol{\alpha}'(t_1)$ es tangente a C, por lo que $\nabla f\ [\boldsymbol{\alpha}(t_1)]$ está en el plano normal a C, como muestra la figura 9.9.

Los dos vectores gradientes ∇g_1 y ∇g_2 son independientes si y sólo si su producto vectorial es no nulo. Este producto viene dado por

$$\nabla g_1 \times \nabla g_2 = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} & \dfrac{\partial g_1}{\partial z} \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} & \dfrac{\partial g_2}{\partial z} \end{vmatrix} = \frac{\partial(g_1, g_2)}{\partial(y, z)}\boldsymbol{i} + \frac{\partial(g_1, g_2)}{\partial(z, x)}\boldsymbol{j} + \frac{\partial(g_1, g_2)}{\partial(x, y)}\boldsymbol{k}.$$

Por consiguiente, la independencia de ∇g_1 *y* ∇g_2 significa que no todos los determinantes jacobianos del segundo miembro son cero. Como ya hemos observado, el método de Lagrange es aplicable siempre que esta condición se satisfaga.

Si ∇g_1 y ∇g_2 son dependientes el método puede fallar. Por ejemplo, supongamos que intentamos la aplicación del método de Lagrange para encontrar los valores extremos de $f(x, y, z) = x^2 + y^2$ en la curva de intersección de las dos superficies $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$ siendo $g_1(x, y, z) = z$ y $g_2(x, y, z) = z^2 - (y - 1)^3$. Las dos superficies, un plano y un cilindro, se cortan a lo largo de la recta C dibujada en la figura 9.10. El problema tiene evidentemente una

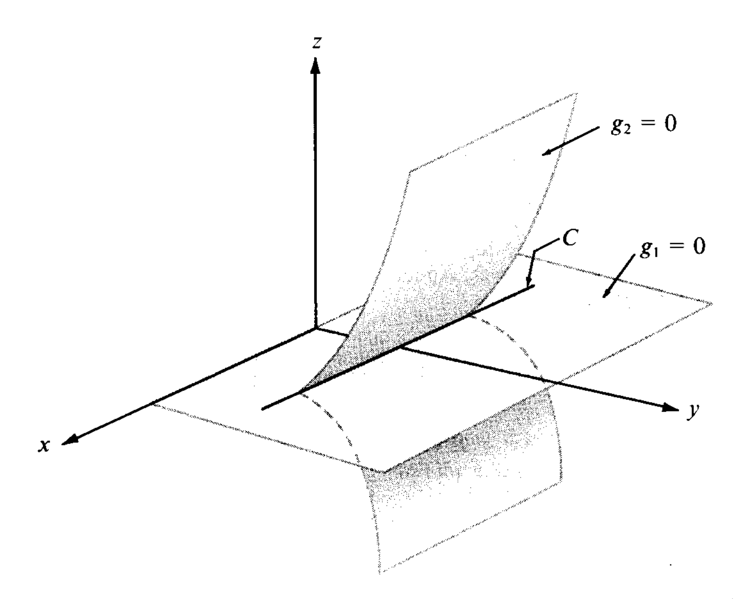


FIGURA 9.10 *Ejemplo en el que el método de Lagrange no es aplicable.*

solución, debido a que $f(x, y, z)$ representa la distancia del punto (x, y, z) al eje z y esta distancia es un mínimo sobre C cuando el punto es el $(0, 1, 0)$. Sin embargo, en este punto los vectores gradientes son $\nabla g_1 = \boldsymbol{k}$, $\nabla g_2 = \boldsymbol{O}$, y $\nabla f = 2\boldsymbol{j}$, y está claro que no existen escalares λ_1 y λ_2 que satisfagan la ecuación (9.41).

9.15 Ejercicios

1. Hallar los valores extremos de $z = xy$ con la condición $x + y = 1$.
2. Hallar las distancias máxima y mínima desde el origen a la curva $5x^2 + 6xy + 5y^2 = 8$.
3. Supongamos que a y b son números positivos fijos.
   a) Hallar los valores extremos de $z = x/a + y/b$ con la condición $x^2 + y^2 = 1$.
   b) Hallar los valores extremos de $z = x^2 + y^2$ con la condición $x/a + y/b = 1$.
   En cada caso, interpretar geométricamente el problema.
4. Hallar los valores extremos de $z = \cos^2 x + \cos^2 y$ con la condición $x - y = \pi/4$.
5. Hallar los valores extremos del campo escalar $f(x, y, z) = x - 2y + 2z$ en la esfera $x^2 + y^2 + z^2 = 1$.

6. Hallar los puntos de la superficie $z^2 - xy = 1$ más próximos al origen.
7. Hallar la mínima distancia desde el punto $(1, 0)$ a la parábola $y^2 = 4x$.
8. Hallar los puntos de la curva de intersección de las dos superficies

$$x^2 - xy + y^2 - z^2 = 1 \qquad \text{y} \qquad x^2 + y^2 = 1$$

que están más próximos al origen.
9. Si a, b y c son números positivos, hallar el valor máximo de $f(x, y, z) = x^a y^b z^c$ con la condición $x + y + z = 1$.
10. Hallar el volumen mínimo limitado por los planos $x = 0$, $y = 0$, $z = 0$, y un plano que sea tangente al elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

en un punto del octante $x > 0$, $y > 0$, $z > 0$.
11. Hallar el máximo de $\log x + \log y + 3 \log z$ en la porción de la esfera $x^2 + y^2 + z^2 = 5r^2$ en la que $x > 0$, $y > 0$, $z > 0$. Aplicar el resultado para demostrar que para números reales positivos a, b, c tenemos

$$abc^3 \leq 27\left(\frac{a + b + c}{5}\right)^5.$$

12. Dada la sección cónica $Ax^2 + 2Bxy + Cy^2 = 1$, siendo $A > 0$ y $B^2 < AC$. Representemos con m y M las distancias mínima y máxima desde el origen a los puntos de la cónica. Demostrar que

$$M^2 = \frac{A + C + \sqrt{(A - C)^2 + 4B^2}}{2(AC - B^2)}$$

y hallar una fórmula análoga para m^2.
13. Aplicar el método de los multiplicadores de Lagrange para hallar las distancias máxima y mínima de un punto de la elipse $x^2 + 4y^2 = 4$ a la recta $x + y = 4$.
14. La sección de un canal es un trapecio isósceles. Si los lados iguales de ese trapecio miden c metros, ¿cuál debe ser el ángulo que forman éstos con el fondo (base menor del trapecio) si queremos que el área de la sección sea máxima?

9.16 Teorema del valor extremo para campos escalares continuos

El teorema del valor extremo para funciones reales continuas en un intervalo acotado y cerrado puede extenderse a campos escalares. Consideremos campos escalares continuos en un intervalo n-dimensional cerrado. Tal intervalo se define como el producto cartesiano de n intervalos uni-dimensionales cerrados. Si $\boldsymbol{a} = (a_1, \ldots, a_n)$ y $\boldsymbol{b} = (b_1, \ldots, b_n)$ escribimos

$$[\boldsymbol{a}, \boldsymbol{b}] = [a_1, b_1] \times \cdots \times [a_n, b_n] = \{(x_1, \ldots, x_n) \mid x_1 \in [a_1, b_1], \ldots, x_n \in [a_n, b_n]\}.$$

Por ejemplo, cuando $n = 2$ el producto cartesiano $[\mathbf{a}, \mathbf{b}]$ es un rectángulo.

La demostración del teorema del valor extremo es paralela a la demostración dada en el volumen I para el caso unidimensional. Demostramos primero que la continuidad de f implica la acotación, luego probamos que f alcanza efectivamente sus valores máximo y mínimo en $[\mathbf{a}, \mathbf{b}]$.

TEOREMA 9.8. TEOREMA DE ACOTACIÓN PARA CAMPOS ESCALARES CONTINUOS. *Si f es un campo escalar continuo en cada punto de un intervalo cerrado $[\mathbf{a}, \mathbf{b}]$ de $\mathbf{R}^n$, entonces f es acotada en $[\mathbf{a}, \mathbf{b}]$. Esto es, existe un número $C \geqslant 0$ tal que $|f(\mathbf{x})| \leqslant C$ para todo $\mathbf{x}$ de $[\mathbf{a}, \mathbf{b}]$.*

Demostración. Razonemos por reducción al absurdo, utilizando el método de la bisección sucesiva. La figura 9.11 representa el método para el caso $n = 2$.

Supongamos que f no esté acotada en $[\mathbf{a}, \mathbf{b}]$. Pongamos $I^{(1)} = [\mathbf{a}, \mathbf{b}]$ e $I_k^{(1)} = [a_k, b_k]$, así que

$$I^{(1)} = I_1^{(1)} \times \cdots \times I_n^{(1)} .$$

Dividamos en dos partes iguales cada intervalo unidimensional $I_k^{(1)}$ formando dos subintervalos, la mitad izquierda $I_{k,1}^{(1)}$ y la mitad derecha $I_{k,2}^{(1)}$. Consideremos ahora todos los productos cartesianos posibles de la forma

$$I_{1,j_1}^{(1)} \times \cdots \times I_{n,j_n}^{(1)} ,$$

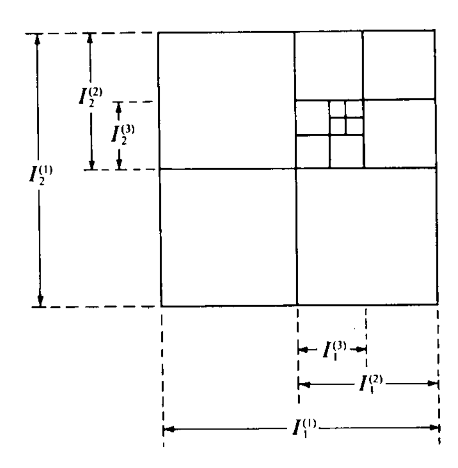


FIGURA 9.11 *Representación en el plano del método de la bisección sucesiva.*

en donde cada $j_i = 1$ ó 2. Existen exactamente 2^n productos de este tipo. Cada uno de ellos es un subintervalo n-dimensional de $[\mathbf{a}, \mathbf{b}]$, y su reunión es igual a $[\mathbf{a}, \mathbf{b}]$. La función f no está acotada en *uno por lo menos* de esos subintervalos (si estuviera acotada en cada uno de ellos también estaría en $[\mathbf{a}, \mathbf{b}]$). Designemos por $I^{(2)}$ uno de ellos que expresamos del modo siguiente

$$I^{(2)} = I_1^{(2)} \times \cdots \times I_n^{(2)},$$

en donde cada $I_k^{(2)}$ es uno de los subintervalos unidimensionales de $I_k^{(1)}$, de longitud $\frac{1}{2}(b_k - a_k)$.

Seguidamente hacemos con $I^{(2)}$ lo mismo que con $I^{(1)}$, bisecamos cada componente unidimensional $I_k^{(2)}$ y llegamos a un intervalo n-dimensional $I^{(3)}$ en el que f no está acotada. Prosiguiendo este proceso, obtenemos un conjunto infinito de intervalos n-dimensionales

$$I^{(1)}, I^{(2)}, \ldots, \text{tales que } I^{(m+1)} \subseteq I^{(m)},$$

en cada uno de los cuales f no está acotada. El intervalo m-ésimo $I^{(m)}$ puede expresarse en la forma

$$I^{(m)} = I_1^{(m)} \times \cdots \times I_n^{(m)}.$$

Como quiera que cada intervalo unidimensional $I_k^{(m)}$ se obtiene mediante $m - 1$ bisecciones sucesivas de $[a_k, b_k]$, si escribimos $I_k^{(m)} = [a_k^{(m)}, b_k^{(m)}]$ tenemos

$$(9.42) \qquad b_k^{(m)} - a_k^{(m)} = \frac{b_k - a_k}{2^{m-1}}, \quad \text{para} \quad k = 1, 2, \ldots, n.$$

Para cada k fija, el extremo superior de todos los extremos izquierdos $a_k^{(m)}$ ($m = 1, 2, \ldots$) debe ser igual al extremo inferior de todos los extremos derechos $b_k^{(m)}$ ($m = 1, 2, \ldots$), y designamos su valor común con t_k. El punto $\mathbf{t} = (t_1, \ldots, t_n)$ está contenido en $[\mathbf{a}, \mathbf{b}]$. En virtud de la continuidad de f en $\mathbf{t}$ existe una n-bola $B(\mathbf{t}; r)$ en la que tenemos

$$|f(\mathbf{x}) - f(\mathbf{t})| \leq 1 \text{ para todo } \mathbf{x} \text{ de } B(\mathbf{t}; r) \cap [\mathbf{a}, \mathbf{b}].$$

Esta desigualdad implica

$$|f(\mathbf{x})| < 1 + |f(\mathbf{t})| \text{ para todo } \mathbf{x} \text{ de } B(\mathbf{t}; r) \cap [\mathbf{a}, \mathbf{b}],$$

con lo que f no está acotada en el conjunto $B(\mathbf{t}; r) \cap [\mathbf{a}, \mathbf{b}]$. Pero este conjunto

contiene todo el intervalo $I^{(m)}$ cuando m es lo bastante grande para que cada uno de los n números (9.42) sea menor que $r/\sqrt{n}$. Por consiguiente para ese valor de m la función f no está acotada en $I^{(m)}$, lo que está en contradicción con el hecho de que f está acotada en $I^{(m)}$. Con ello se completa la demostración.

Si f está acotada en $[\boldsymbol{a}, \boldsymbol{b}]$, el conjunto de todos los valores $f(\boldsymbol{x})$ es un conjunto de números reales acotado superior e inferiormente. Por consiguiente ese conjunto tiene un extremo superior y un extremo inferior que designamos con $\sup f$ e $\inf f$, respectivamente. Esto es, escribimos

$$\sup f = \sup \{f(\boldsymbol{x}) \mid \boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]\}, \qquad \inf f = \inf \{f(\boldsymbol{x}) \mid \boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]\}.$$

Ahora vamos a demostrar que una función continua alcanza los valores $\inf f$ y $\sup f$ en $[\boldsymbol{a}, \boldsymbol{b}]$.

TEOREMA 9.9. TEOREMA DEL VALOR EXTREMO PARA CAMPOS ESCALARES. *Si f es continua en un intervalo cerrado $[\boldsymbol{a}, \boldsymbol{b}]$ de $\mathbf{R}^n$, entonces existen puntos $\boldsymbol{c}$ y $\boldsymbol{d}$ en $[\boldsymbol{a}, \boldsymbol{b}]$ tales que*

$$f(\boldsymbol{c}) = \sup f \qquad \text{y} \qquad f(\boldsymbol{d}) = \inf f.$$

Demostración. Basta demostrar que f alcanza su extremo superior en $[\boldsymbol{a}, \boldsymbol{b}]$. El resultado para el extremo inferior se deduce como una consecuencia debido a que el extremo inferior de f es el extremo superior de $-f$.

Pongamos $M = \sup f$. Supondremos que en $[\boldsymbol{a}, \boldsymbol{b}]$ no existe ningún $\boldsymbol{x}$ para el que $f(\boldsymbol{x}) = M$ y obtendremos una contradicción. Pongamos $g(\boldsymbol{x}) = M - f(\boldsymbol{x})$. Para todo $\boldsymbol{x}$ de $[\boldsymbol{a}, \boldsymbol{b}]$ es entonces $g(\boldsymbol{x}) > 0$ de modo que la función recíproca $1/g$ es continua en $[\boldsymbol{a}, \boldsymbol{b}]$. Según el teorema de acotación, $1/g$ está acotada en $[\boldsymbol{a}, \boldsymbol{b}]$, sea $1/g(\boldsymbol{x}) < C$ para todo $\boldsymbol{x}$ de $[\boldsymbol{a}, \boldsymbol{b}]$, siendo $C > 0$. Esto implica que $M - f(\boldsymbol{x}) > 1/C$, con lo que $f(\boldsymbol{x}) < M - 1/C$ para todo $\boldsymbol{x}$ de $[\boldsymbol{a}, \boldsymbol{b}]$. Esto contradice el hecho de que M es la menor cota superior de f en $[\boldsymbol{a}, \boldsymbol{b}]$. Luego $f(\boldsymbol{x}) = M$ para un $\boldsymbol{x}$ por lo menos de $[\boldsymbol{a}, \boldsymbol{b}]$.

9.17 Teorema de la continuidad uniforme para campos escalares continuos

Sea f continua en un intervalo cerrado acotado $[\boldsymbol{a}, \boldsymbol{b}]$ en $\mathbf{R}^n$, y designemos con $M(f)$ y $m(f)$, respectivamente, los valores máximo y mínimo de f en $[\boldsymbol{a}, \boldsymbol{b}]$. La diferencia

$$M(f) - m(f)$$

se llama *oscilación* de f en $[\boldsymbol{a}, \boldsymbol{b}]$. Como en el caso unidimensional tenemos un

teorema de la oscilación uniforme para funciones continuas, que nos dice que en el intervalo $[\mathbf{a}, \mathbf{b}]$ puede efectuarse una partición de modo que la oscilación de f en cada subintervalo sea tan pequeña como se quiera.

Pongamos $[\mathbf{a}, \mathbf{b}] = [a_1, b_1] \times \ldots \times [a_n, b_n]$, y sea P_k una partición del intervalo $[a_k, b_k]$. Esto es, P_k es un conjunto de puntos

$$P_k = \{x_0, x_1, \ldots, x_{r-1}, x_r\}$$

tales que $a_k = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_{r-1} \leqslant x_r = b_k$. El producto cartesiano

$$P = P_1 \times \cdots \times P_n$$

se denomina partición del intervalo $[\mathbf{a}, \mathbf{b}]$. El teorema de la oscilación uniforme se suele llamar de la continuidad uniforme, y toma la siguiente forma.

TEOREMA 9.10. *Si f es un campo escalar continuo en un intervalo cerrado $[\mathbf{a}, \mathbf{b}]$ de $\mathbf{R}^n$, entonces para todo $\epsilon > 0$ existe una partición de $[\mathbf{a}, \mathbf{b}]$ en un número finito de subintervalos tales que la oscilación de f en todo subintervalo es menor que ϵ.*

Demostración. La demostración es del todo análoga al caso unidimensional, así que únicamente esbozamos los principales pasos. Razonamos por reducción al absurdo, con el método de la bisección sucesiva. Suponemos que el teorema es falso; esto es, suponemos que para un cierto ϵ_0 en el intervalo $[\mathbf{a}, \mathbf{b}]$ no puede efectuarse una partición en un número finito de subintervalos en cada uno de los cuales la oscilación de f es menor que ϵ_0. Por bisección sucesiva obtenemos un conjunto infinito de subintervalos $I^{(1)}, I^{(2)}, \ldots$, en cada uno de los cuales la oscilación de f es por lo menos ϵ_0. Considerando la menor cota superior de los extremos izquierdos de los intervalos componentes de $I^{(1)}, I^{(2)}, \ldots$, obtenemos un punto $\mathbf{t}$ de $[\mathbf{a}, \mathbf{b}]$ perteneciente a todos estos subintervalos. En virtud de la continuidad de f en $\mathbf{t}$ existe una n-bola $B(\mathbf{t}; r)$ tal que la oscilación de f es menor que $\frac{1}{2}\epsilon_0$ en $B(\mathbf{t}; r) \cap [\mathbf{a}, \mathbf{b}]$. Pero, cuando m es suficientemente grande, el intervalo $I^{(m)}$ está contenido en el conjunto $B(\mathbf{t}; r) \cap [\mathbf{a}, \mathbf{b}]$, por lo que la oscilación de f no supera *a* $\frac{1}{2}\epsilon_0$ en $I^{(m)}$, lo que está en contradicción con el hecho de que la oscilación de f es por lo menos ϵ_0 en $I^{(m)}$.

10

INTEGRALES DE LÍNEA

10.1 Introducción

En el volumen I estudiamos la integral $\int_a^b f(x)\,dx$, primero para funciones reales definidas y acotadas en intervalos finitos, y luego para funciones no acotadas e intervalos infinitos. Posteriormente el concepto se extendió a funciones vectoriales y, en el capítulo 7 del volumen II, a funciones matriciales.

Este capítulo extiende la noción de integral en otra dirección. El intervalo $[a, b]$ se reemplaza por una curva en el espacio n-dimensional definida por una función vectorial $\boldsymbol{\alpha}$, y el integrando es un campo vectorial $\boldsymbol{f}$ definido y acotado en esa curva. La integral que resulta se llama *integral de línea, integral curvilínea* o *integral de contorno,* y se emplea para ella la notación $\int \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ o algún otro símbolo parecido. El punto se usa precisamente para sugerir el producto interior de dos vectores. La curva se llama *camino de integración.*

Las integrales de línea son de capital importancia en Matemática pura y aplicada. Se presentan al estudiar el trabajo, la energía potencial, el flujo de calor, el cambio en la entropía, la circulación de un fluido, y otras cuestiones físicas en las que se estudia el comportamiento de un campo escalar o vectorial a lo largo de una curva.

10.2 Caminos e integrales de línea

Antes de definir las integrales de línea recordemos la definición de curva dada en el volumen I. Sea $\boldsymbol{\alpha}$ una función vectorial definida en un intervalo cerrado finito $J = [a, b]$. Cuando t va tomando los valores de J, la función $\boldsymbol{\alpha}(t)$ describe un conjunto de puntos en el n-espacio llamado *gráfica* de la función. Si $\boldsymbol{\alpha}$ es continua en J la gráfica se llama *curva*; con mayor precisión, es la curva descrita por $\boldsymbol{\alpha}$.

En nuestro estudio de las curvas en el volumen I vimos que funciones dis-

tintas pueden originar el trazado de la misma curva en formas distintas, por ejemplo, en direcciones distintas o con velocidades distintas. Al estudiar las integrales de línea nos interesa no sólo el conjunto de puntos de una curva sino la manera como tal curva ha sido originada, esto es, la función $\boldsymbol{\alpha}$. Una tal función se llamará *camino continuo.*

DEFINICIÓN. *Sea $J = [a, b]$ un intervalo cerrado finito de $\mathbf{R}^1$. Una función $\boldsymbol{\alpha}: J \rightarrow \mathbf{R}^n$ continua en J se llama camino continuo en el n-espacio. El camino se llama regular si existe la derivada $\boldsymbol{\alpha}'$ y es continua en el intervalo abierto (a, b). El camino se llama regular a trozos si el intervalo $[a, b]$ puede descomponerse en un número finito de subintervalos en cada uno de los cuales el camino es regular.*

La figura 10.1 muestra la gráfica de un camino regular a trozos. En este ejemplo la curva tiene recta tangente en todos los puntos excepto en un número finito de ellos. Esos puntos excepcionales subdividen la curva en arcos, a lo largo de cada uno de los cuales la recta tangente va cambiando de posición con continuidad.

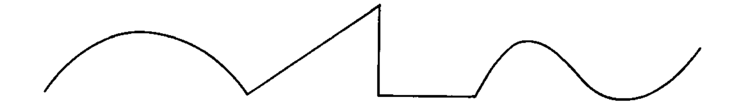


FIGURA 10.1 *Gráfica de un camino regular a trozos en el plano.*

DEFINICIÓN DE INTEGRAL DE LÍNEA. *Sea $\boldsymbol{\alpha}$ un camino regular a trozos en el n-espacio definido en un intervalo $[a, b]$, y sea $\boldsymbol{f}$ un campo vectorial definido y acotado sobre la gráfica de $\boldsymbol{\alpha}$. La integral de línea de $\boldsymbol{f}$ a lo largo de $\boldsymbol{\alpha}$ se representa con el símbolo $\int \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ y se define por*

$$(10.1) \qquad \int \boldsymbol{f} \cdot d\boldsymbol{\alpha} = \int_a^b \boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t)\, dt,$$

siempre que la integral del segundo miembro exista, bien como integral propia o integral impropia.

Observación: En muchos ejemplos que en la práctica se presentan el producto interior $\boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t)$ está acotado en $[a, b]$ y es continuo excepto, acaso, en un número finito de puntos, en cuyo caso la integral existe como integral propia.

10.3 Otras notaciones para las integrales de línea

Si C representa la gráfica de $\boldsymbol{\alpha}$, la integral de línea $\int \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ también se repre-

senta por $\int_C \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ y se llama *integral de* $\boldsymbol{f}$ *a lo largo de* C.

Si $\boldsymbol{a} = \boldsymbol{\alpha}(a)$ y $\boldsymbol{b} = \boldsymbol{\alpha}(b)$ representan los puntos extremos de C, a veces la integral de línea se expresa poniendo $\int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{f}$ o $\int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ y se denomina integral de línea de $\boldsymbol{f}$ desde $\boldsymbol{a}$ a $\boldsymbol{b}$ a lo largo de $\boldsymbol{\alpha}$. Cuando se use la notación $\int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{f}$ deberá tenerse en cuenta que la integral depende no solamente de los extremos $\boldsymbol{a}$ y $\boldsymbol{b}$ sino también del camino $\boldsymbol{\alpha}$ que los une.

Cuando $\boldsymbol{a} = \boldsymbol{b}$ el camino se llama *cerrado*. A menudo el símbolo $\oint$ se usa para indicar la integración a lo largo de un camino cerrado.

Cuando $\boldsymbol{f}$ y $\boldsymbol{\alpha}$ se expresan en función de sus componentes, a saber

$$\boldsymbol{f} = (f_1, \ldots, f_n) \qquad \text{y} \qquad \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n),$$

la integral del segundo miembro de (10.1) se convierte en una suma de integrales,

$$\sum_{k=1}^{n} \int_a^b f_k[\boldsymbol{\alpha}(t)]\alpha_k'(t)\, dt .$$

En este caso la integral de línea también se pone en la forma $\int f_1\, d\alpha_1 + \ldots + f_n\, d\alpha_n$.

En el caso bi-dimensional ordinariamente el camino $\boldsymbol{\alpha}$ se define con un par de ecuaciones paramétricas,

$$x = \alpha_1(t), \qquad y = \alpha_2(t),$$

y la integral de línea $\int_C \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ se escribe en la forma $\int_C f_1\, dx + f_2\, dy$, o bien $\int_C f_1(x, y)\, dx + f_2(x, y)\, dy$.

En tres dimensiones se utilizan tres ecuaciones paramétricas

$$x = \alpha_1(t), \qquad y = \alpha_2(t), \qquad z = \alpha_3(t)$$

y ponemos la integral $\int_C \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ en la forma $\int_C f_1\, dx + f_2\, dy + f_3\, dz$, o bien

$$\int_C f_1(x, y, z)\, dx + f_2(x, y, z)\, dy + f_3(x, y, z)\, dz .$$

EJEMPLO. Sea $\boldsymbol{f}$ un campo vectorial de dos dimensiones dado por

$$\boldsymbol{f}(x, y) = \sqrt{y}\, \boldsymbol{i} + (x^3 + y)\boldsymbol{j}$$

para todo (x, y) con $y \geq 0$. Calcular la integral de línea de $\boldsymbol{f}$ desde $(0, 0)$ a $(1, 1)$ a lo largo de cada uno de los siguientes caminos:

a) la recta de ecuaciones paramétricas $x = t$, $y = t$, $0 \leq t \leq 1$;
b) el camino de ecuaciones paramétricas $x = t^2$, $y = t^3$, $0 \leq t \leq 1$.

Solución. Para el camino de la parte a) tomamos $\boldsymbol{\alpha}(t) = t\boldsymbol{i} + t\boldsymbol{j}$. Entonces $\boldsymbol{\alpha}'(t) = \boldsymbol{i} + \boldsymbol{j}$ y $\boldsymbol{f}[\boldsymbol{\alpha}(t)] = \sqrt{t}\,\boldsymbol{i} + (t^3 + t)\boldsymbol{j}$. Por consiguiente el producto interior de $\boldsymbol{f}[\boldsymbol{\alpha}(t)]$ por $\boldsymbol{\alpha}'(t)$ es igual a $\sqrt{t} + t^3 + t$ y encontramos

$$\int_{(0,0)}^{(1,1)} \boldsymbol{f} \cdot d\boldsymbol{\alpha} = \int_0^1 (\sqrt{t} + t^3 + t)\, dt = \frac{17}{12}.$$

Para el camino de la parte b) tomamos $\boldsymbol{\alpha}(t) = t^2\boldsymbol{i} + t^3\boldsymbol{j}$. Entonces $\boldsymbol{\alpha}'(t) = 2t\boldsymbol{i} + 3t^2\boldsymbol{j}$ y $\boldsymbol{f}[\boldsymbol{\alpha}(t)] = t^{3/2}\boldsymbol{i} + (t^6 + t^3)\boldsymbol{j}$. Por consiguiente

$$\boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t) = 2t^{5/2} + 3t^8 + 3t^5,$$

así que

$$\int_{(0,0)}^{(1,1)} \boldsymbol{f} \cdot d\boldsymbol{\alpha} = \int_0^1 (2t^{5/2} + 3t^8 + 3t^5)\, dt = \frac{59}{42}.$$

Estos ejemplos ponen de manifiesto que la integral desde un punto a otro puede depender del camino que los une.

Calculemos ahora la parte b) una vez más, utilizando la misma curva pero con representación paramétrica distinta. Dicha curva puede representarse con la función

$$\boldsymbol{\beta}(t) = t\boldsymbol{i} + t^{3/2}\boldsymbol{j}, \qquad \text{donde} \quad 0 \leq t \leq 1.$$

Esto nos lleva a la relación

$$\boldsymbol{f}[\boldsymbol{\beta}(t)] \cdot \boldsymbol{\beta}'(t) = (t^{3/4}\boldsymbol{i} + (t^3 + t^{3/2})\boldsymbol{j}) \cdot (\boldsymbol{i} + \tfrac{3}{2}t^{1/2}\boldsymbol{j}) = t^{3/4} + \tfrac{3}{2}t^{7/2} + \tfrac{3}{2}t^2,$$

cuya integral desde 0 a 1 es 59/42, como antes. Esto nos hace ver que el valor de la integral es independiente de la representación paramétrica utilizada para la curva. Esta es una propiedad general de las integrales de línea que se demuestra en la sección siguiente.

10.4 Propiedades fundamentales de las integrales de línea

Puesto que las integrales de línea se definen en función de integrales ordinarias, no debe sorprender que aquéllas gocen de muchas de las propiedades de éstas. Por ejemplo, tienen la *propiedad de linealidad* respecto al integrando,

$$\int (a\boldsymbol{f} + b\boldsymbol{g}) \cdot d\boldsymbol{\alpha} = a \int \boldsymbol{f} \cdot d\boldsymbol{\alpha} + b \int \boldsymbol{g} \cdot d\boldsymbol{\alpha},$$

y la *propiedad aditiva* respecto al camino de integración:

$$\int_C f \cdot d\boldsymbol{\alpha} = \int_{C_1} f \cdot d\boldsymbol{\alpha} + \int_{C_2} f \cdot d\boldsymbol{\alpha},$$

donde las dos curvas C_1 y C_2 forman la curva C. Esto es, C es la gráfica de una función $\boldsymbol{\alpha}$ definida en un intervalo $[a, b]$, y las curvas C_1 y C_2 son las representaciones gráficas de $\boldsymbol{\alpha}(t)$ al variar t en los subintervalos $[a, c]$ y $[c, b]$ respectivamente, para un c que cumple $a < c < b$. Las demostraciones de estas propiedades son consecuencia inmediata de la definición de la integral de línea; y las dejamos como ejercicios para el lector.

Seguidamente examinamos el comportamiento de las integrales de línea al efectuar un cambio de parámetro. Sea $\boldsymbol{\alpha}$ un camino continuo definido en un intervalo $[a, b]$, sea u una función real derivable, de modo que u' nunca sea cero en un intervalo $[c, d]$, y tal que el recorrido de u sea $[a, b]$. Entonces la función $\boldsymbol{\beta}$ definida en $[c, d]$ por la ecuación

$$\boldsymbol{\beta}(t) = \boldsymbol{\alpha}[u(t)]$$

es un camino continuo que tiene la misma gráfica que $\boldsymbol{\alpha}$. Dos caminos $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ así relacionados se llaman *equivalentes*. Se dice que proporcionan distintas representaciones paramétricas de la misma curva. Se dice que la función u define un cambio de parámetro.

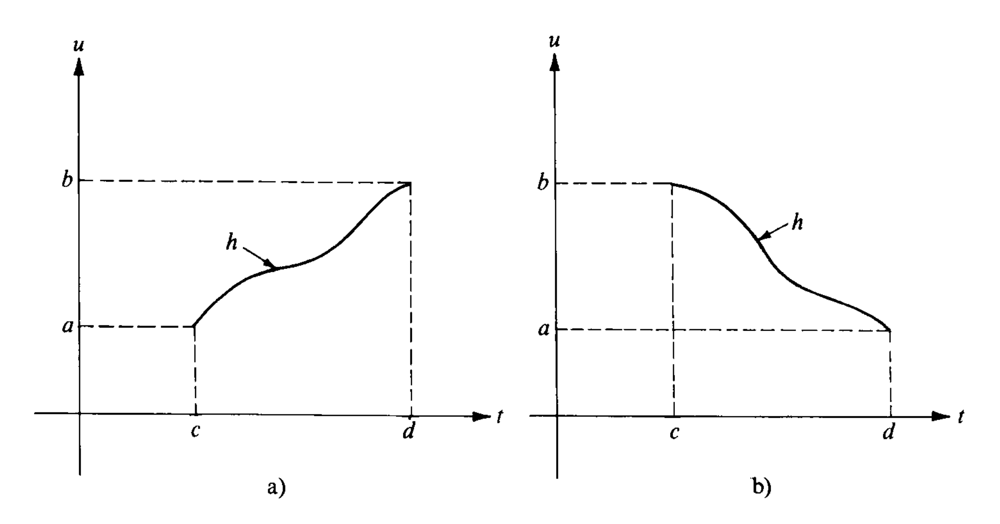


FIGURA 10.2 *Cambio de parámetro definido por* $u = h(t)$. *En* a), *la función* h *conserva la orientación. En* b), *la función* h *invierte la orientación.*

Sea C la gráfica común de los dos caminos equivalentes $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$. Si la derivada de u es siempre positiva en $[c, d]$ la función u es creciente y decimos que los dos caminos $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ originan C en la *misma dirección.* Si la derivada de u es siempre negativa decimos que $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ originan C en *direcciones opuestas.* En el primer caso se dice que *u conserva la orientación;* y en el segundo caso que *u invierte la orientación.* En la figura 10.2 se muestra un ejemplo.

El teorema siguiente demuestra que una integral de línea no varía al efectuar un cambio de parámetro que conserva la orientación; cambia de signo si el cambio de parámetro invierte la orientación. Se supone que existen las dos integrales $\int \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ e $\int \boldsymbol{f} \cdot d\boldsymbol{\beta}$

TEOREMA 10.1. COMPORTAMIENTO DE UNA INTEGRAL DE LINEA FRENTE A UN CAMBIO DE PARÁMETRO. *Si $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ son dos caminos equivalentes regulares a trozos, entonces se tiene*

$$\int_C \boldsymbol{f} \cdot d\boldsymbol{\alpha} = \int_C \boldsymbol{f} \cdot d\boldsymbol{\beta}$$

si $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ originan C en la misma dirección; y

$$\int_C \boldsymbol{f} \cdot d\boldsymbol{\alpha} = -\int_C \boldsymbol{f} \cdot d\boldsymbol{\beta}$$

si $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ originan C en direcciones opuestas.

Demostración. Basta demostrar el teorema para caminos regulares; luego se aplica la propiedad aditiva con respecto al camino de integración para deducir el resultado para caminos regulares a trozos.

La demostración es una simple aplicación de la regla de la cadena. Los caminos $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ están ligados por una relación de la forma $\boldsymbol{\beta}(t) = \boldsymbol{\alpha}[u(t)]$, estando u definida en un intervalo $[c, d]$ y $\boldsymbol{\alpha}$ en un intervalo $[a, b]$. De la regla de la cadena resulta

$$\boldsymbol{\beta}'(t) = \boldsymbol{\alpha}'[u(t)]u'(t).$$

Por consiguiente encontramos

$$\int_C \boldsymbol{f} \cdot d\boldsymbol{\beta} = \int_c^d \boldsymbol{f}[\boldsymbol{\beta}(t)] \cdot \boldsymbol{\beta}'(t)\, dt = \int_c^d \boldsymbol{f}(\boldsymbol{\alpha}[u(t)]) \cdot \boldsymbol{\alpha}'[u(t)]u'(t)\, dt.$$

En la última integral hacemos la sustitución $v = u(t)$, $dv = u'(t)\, dt$ y se obtiene

$$\int_C \boldsymbol{f} \cdot d\boldsymbol{\beta} = \int_{u(c)}^{u(d)} \boldsymbol{f}(\boldsymbol{\alpha}(v)) \cdot \boldsymbol{\alpha}'(v)\, dv = \pm\int_a^b \boldsymbol{f}(\boldsymbol{\alpha}(v)) \cdot \boldsymbol{\alpha}'(v)\, dv = \pm\int_C \boldsymbol{f} \cdot d\boldsymbol{\alpha},$$

en donde se utiliza el signo + si $a = u(c)$ y $b = u(d)$, y el signo − si $a = u(d)$ y $b = u(c)$. El primer caso se presenta si $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ originan C en la misma dirección, el segundo si originan C en direcciones opuestas.

10.5 Ejercicios

En cada uno de los ejercicios 1 al 8 calcular la integral de línea del campo vectorial f a lo largo del camino que se indica.

1. $f(x, y) = (x^2 - 2xy)\boldsymbol{i} + (y^2 - 2xy)\boldsymbol{j}$, a lo largo de la parábola $y = x^2$ desde $(-1, 1)$ a $(1, 1)$.
2. $f(x, y) = (2a - y)\boldsymbol{i} + x\boldsymbol{j}$, a lo largo del camino descrito por $\boldsymbol{\alpha}(t) = a(t - \operatorname{sen} t)\boldsymbol{i} + a(1 - \cos t)\boldsymbol{j}, 0 \leq t \leq 2\pi$.
3. $f(x, y, z) = (y^2 - z^2)\boldsymbol{i} + 2yz\boldsymbol{j} - x^2\boldsymbol{k}$, a lo largo del camino descrito por $\boldsymbol{\alpha}(t) = t\boldsymbol{i} + t^2\boldsymbol{j} + t^3\boldsymbol{k}$, $0 \leq t \leq 1$.
4. $f(x, y) = (x^2 + y^2)\boldsymbol{i} + (x^2 - y^2)\boldsymbol{j}$, a lo largo de la curva $y = 1 - |1 - x|$, desde $(0,0)$ a $(2, 0)$.
5. $f(x, y) = (x + y)\boldsymbol{i} + (x - y)\boldsymbol{j}$, alrededor de la elipse $b^2x^2 + a^2y^2 = a^2b^2$ en sentido contrario al de las agujas del reloj.
6. $f(x, y, z) = 2xy\boldsymbol{i} + (x^2 + z)\boldsymbol{j} + y\boldsymbol{k}$, desde $(1, 0, 2)$ a $(3, 4, 1)$ a lo largo de un segmento de recta.
7. $f(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + (xz - y)\boldsymbol{k}$, desde $(0, 0, 0)$ a $(1, 2, 4)$ a lo largo de un segmento rectilíneo.
8. $f(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + (xz - y)\boldsymbol{k}$, a lo largo del camino dado por $\boldsymbol{\alpha}(t) = t^2\boldsymbol{i} + 2t\boldsymbol{j} + 4t^3\boldsymbol{k}$, $0 \leq t \leq 1$.

En cada uno de los ejercicios 9 al 12, calcular el valor de la integral de línea dada.

9. $\int_C (x^2 - 2xy)\, dx + (y^2 - 2xy)\, dy$ siendo C el arco de parábola $y = x^2$ que une los puntos $(-2, 4)$ y $(1, 1)$.
10. $\displaystyle\int_C \frac{(x + y)\, dx - (x - y)\, dy}{x^2 + y^2}$ donde C es la circunferencia $x^2 + y^2 = a^2$, recorrida en sentido contrario al de las agujas del reloj.
11. $\displaystyle\int_C \frac{dx + dy}{|x| + |y|}$ donde C es el contorno del cuadrado de vértices $(1, 0)$, $(0, 1)$, $(-1, 0)$ y $(0, -1)$, recorrido en sentido contrario al de las agujas del reloj.
12. $\int_C y\, dx + z\, dy + x\, dz$, donde
 a) C es la curva de intersección de las dos superficies $x + y = 2$ y $x^2 + y^2 + z^2 = 2(x + y)$. La curva es recorrida de tal modo que mirando desde el origen el sentido es el de las agujas del reloj.
 b) C es la intersección de las dos superficies $z = xy$ y $x^2 + y^2 = 1$, recorrida en sentido, que visto desde encima del plano xy, es el contrario al de las agujas del reloj.

10.6 El concepto de trabajo como integral de línea

Consideremos una partícula que se mueve a lo largo de una curva bajo la

acción de un campo de fuerzas f. Si la curva es la gráfica de un camino $\boldsymbol{\alpha}$, regular a trozos, el *trabajo* realizado por f se define por la integral de línea $\int \boldsymbol{f} \cdot d\boldsymbol{\alpha}$. Los ejemplos siguientes ponen de manifiesto algunas de las propiedades fundamentales del trabajo.

EJEMPLO 1. *Trabajo realizado por una fuerza constante.* Si $\boldsymbol{f}$ es una fuerza constante, a saber $\boldsymbol{f} = \boldsymbol{c}$, puede demostrarse que el trabajo realizado por $\boldsymbol{f}$ al mover una partícula desde un punto $\boldsymbol{a}$ a un punto $\boldsymbol{b}$ a lo largo de cualquier camino regular a trozos que una $\boldsymbol{a}$ y $\boldsymbol{b}$ es $\boldsymbol{c} \cdot (\boldsymbol{b} - \boldsymbol{a})$ producto de la fuerza por el desplazamiento $\boldsymbol{b} - \boldsymbol{a}$. Lo demostraremos en un caso particular.

Sea $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ un camino que una $\boldsymbol{a}$ y $\boldsymbol{b}$, a saber $\boldsymbol{\alpha}(a) = \boldsymbol{a}$ y $\boldsymbol{\alpha}(b) = \boldsymbol{b}$, y escribamos $\boldsymbol{c} = (c_1, \ldots, c_n)$. Supongamos que $\boldsymbol{\alpha}'$ es continua en $[a, b]$. Entonces el trabajo realizado por $\boldsymbol{f}$ es igual a

$$\int \boldsymbol{f} \cdot d\boldsymbol{\alpha} = \sum_{k=1}^{n} c_k \int_a^b \alpha_k'(t)\, dt = \sum_{k=1}^{n} c_k[\alpha_k(b) - \alpha_k(a)] = \boldsymbol{c} \cdot [\boldsymbol{\alpha}(b) - \boldsymbol{\alpha}(a)] = \boldsymbol{c} \cdot (\boldsymbol{b} - \boldsymbol{a}).$$

Para este campo de fuerzas el trabajo depende solamente de los puntos extremos $\boldsymbol{a}$ y $\boldsymbol{b}$ y no de la curva que los une. No todos los campos de fuerza tienen esta propiedad. Los que la tienen se llaman *conservativos*. El ejemplo de la página 395 es un campo de fuerzas no conservativo. En una sección posterior determinaremos todos los campos de fuerzas conservativos.

EJEMPLO 2. *Principio del trabajo y la energía.* Una partícula de masa m se mueve a lo largo de una curva bajo la acción de un campo de fuerzas $\boldsymbol{f}$. Si la velocidad de la partícula en el instante t es $v(t)$, su energía cinética está definida por $\frac{1}{2} mv^2(t)$. Demostrar que *la variación de la energía cinética en cualquier intervalo de tiempo es igual al trabajo realizado por $\boldsymbol{f}$ durante dicho intervalo de tiempo.*

Solución. Designemos por $\boldsymbol{r}(t)$ la posición de la partícula en el instante t. El trabajo realizado por $\boldsymbol{f}$ durante un intervalo de tiempo $[a, b]$ es $\int_{\boldsymbol{r}(a)}^{\boldsymbol{r}(b)} \boldsymbol{f} \cdot d\boldsymbol{r}$. Queremos demostrar que

$$\int_{\boldsymbol{r}(a)}^{\boldsymbol{r}(b)} \boldsymbol{f} \cdot d\boldsymbol{r} = \tfrac{1}{2}mv^2(b) - \tfrac{1}{2}mv^2(a).$$

Según la segunda ley del movimiento de Newton tenemos

$$\boldsymbol{f}[\boldsymbol{r}(t)] = m\boldsymbol{r}''(t) = m\boldsymbol{v}'(t),$$

donde $\boldsymbol{v}(t)$ designa el vector velocidad en el instante t. La velocidad es la longi-

tud del vector velocidad, $v(t) = \|\boldsymbol{v}(t)\|$. Por consiguiente

$$\boldsymbol{f}[\boldsymbol{r}(t)] \cdot \boldsymbol{r}'(t) = \boldsymbol{f}[\boldsymbol{r}(t)] \cdot \boldsymbol{v}(t) = m\boldsymbol{v}'(t) \cdot \boldsymbol{v}(t) = \tfrac{1}{2}m \frac{d}{dt}(\boldsymbol{v}(t) \cdot \boldsymbol{v}(t)) = \tfrac{1}{2}m \frac{d}{dt}(v^2(t)).$$

Integrando entre a y b obtenemos

$$\int_{\boldsymbol{r}(a)}^{\boldsymbol{r}(b)} \boldsymbol{f} \cdot d\boldsymbol{r} = \int_a^b \boldsymbol{f}[\boldsymbol{r}(t)] \cdot \boldsymbol{r}'(t)\, dt = \tfrac{1}{2}mv^2(t)\Big|_a^b = \tfrac{1}{2}mv^2(b) - \tfrac{1}{2}mv^2(a),$$

como queríamos probar.

10.7 Integrales de línea con respecto a la longitud de arco

Sea $\boldsymbol{\alpha}$ un camino con derivada $\boldsymbol{\alpha}'$ continua en un intervalo $[a, b]$. La gráfica de $\boldsymbol{\alpha}$ es una curva rectificable. En el volumen I se demostró que la correspondiente función longitud de arco s, está dada por la integral

$$s(t) = \int_a^t \|\boldsymbol{\alpha}'(u)\|\, du.$$

La derivada de la longitud de arco tiene por valor

$$s'(t) = \|\boldsymbol{\alpha}'(t)\|.$$

Sea φ un campo escalar definido y acotado en C, la gráfica de $\boldsymbol{\alpha}$. *La integral de línea de φ con respecto a la longitud de arco a lo largo de C* se representa con el símbolo $\int_C \varphi\, ds$ y se define por

$$\int_C \varphi\, ds = \int_a^b \varphi[\boldsymbol{\alpha}(t)]s'(t)\, dt,$$

siempre que exista la integral del segundo miembro.

Consideremos ahora un campo escalar φ dado por $\varphi[\boldsymbol{\alpha}(t)] = \boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{T}(t)$ que es el producto interior de un campo vectorial $\boldsymbol{f}$, definido en C, por el vector tangente unitario $\boldsymbol{T}(t) = (d\boldsymbol{\alpha}/ds)$ En este caso la integral de línea $\int_C \varphi\, ds$ coincide con esta otra $\int_C \boldsymbol{f} \cdot d\boldsymbol{\alpha}$ debido a que

$$\boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t) = \boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \frac{d\boldsymbol{\alpha}}{ds}\frac{ds}{dt} = \boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{T}(t)s'(t) = \varphi[\boldsymbol{\alpha}(t)]s'(t).$$

Cuando $\boldsymbol{f}$ representa una velocidad, el producto interior $\boldsymbol{f} \cdot \boldsymbol{T}$ es el componente

tangencial de la velocidad, y la integral de línea $\int_C \boldsymbol{f} \cdot \boldsymbol{T}\, ds$ es la *integral de flujo* de $\boldsymbol{f}$ a lo largo de C. Cuando C es una curva cerrada la integral de flujo es la *circulación* de $\boldsymbol{f}$ a lo largo de C. Estas denominaciones se usan corrientemente en la teoría del flujo de fluidos.

10.8 Otras aplicaciones de las integrales de línea

Las integrales de línea con respecto a la longitud de arco se presentan también en problemas relativos a la distribución de la masa a lo largo de una curva. Por ejemplo, imaginemos una curva C en el espacio de tres dimensiones como un delgado alambre de densidad variable. Supongamos que la densidad se expresa mediante un campo escalar φ, siendo $\varphi(x, y, z)$ la *masa por unidad de longitud* en el punto (x, y, z) de C. La masa total M del alambre viene entonces definida como la integral de línea de φ con respecto a la longitud de arco:

$$M = \int_C \varphi(x, y, z)\, ds.$$

El *centro de gravedad* se define como el punto $(\bar{x}, \bar{y}, \bar{z})$ cuyas coordenadas están determinadas por las ecuaciones

$$\bar{x}M = \int_C x\varphi(x, y, z)\, ds, \qquad \bar{y}M = \int_C y\varphi(x, y, z)\, ds, \qquad \bar{z}M = \int_C z\varphi(x, y, z)\, ds.$$

Un alambre de densidad constante se llama *uniforme*. En este caso el centro de gravedad también se llama *centroide*.

EJEMPLO 1. Calcular la masa M de un muelle que tiene forma de hélice cuya ecuación vectorial es

$$\boldsymbol{\alpha}(t) = a \cos t\, \boldsymbol{i} + a \operatorname{sen} t\, \boldsymbol{j} + bt\boldsymbol{k}$$

si la densidad en (x, y, z) es $x^2 + y^2 + z^2$.

Solución. La integral para calcular M es

$$M = \int_C (x^2 + y^2 + z^2)\, ds = \int_0^{2\pi} (a^2 \cos^2 t + a^2 \operatorname{sen}^2 t + b^2 t^2) s'(t)\, dt.$$

Puesto que $s'(t) = \|\boldsymbol{\alpha}'(t)\|$ y $\boldsymbol{\alpha}'(t) = -a \operatorname{sen} t\boldsymbol{i} + a \cos t\boldsymbol{j} + b\boldsymbol{k}$, tenemos $s'(t) = \sqrt{a^2 + b^2}$ y por tanto

$$M = \sqrt{a^2 + b^2} \int_0^{2\pi} (a^2 + b^2 t^2)\, dt = \sqrt{a^2 + b^2} \left(2\pi a^2 + \frac{8}{3}\pi^3 b^2\right).$$

En este ejemplo la coordenada $\bar{z}$ del centro de gravedad viene dada por

$$\bar{z}M = \int_C z(x^2 + y^2 + z^2)\, ds = \sqrt{a^2 + b^2} \int_0^{2\pi} bt(a^2 + b^2t^2)\, dt$$
$$= b\sqrt{a^2 + b^2}\,(2\pi^2a^2 + 4\pi^4b^2).$$

La determinación de las coordenadas $\bar{x}$ e $\bar{y}$, se propone como ejercicio 15 en la sección 10.9.

Las integrales de línea se pueden utilizar para definir el momento de inercia de un alambre o hilo con respecto a un eje. Si $\delta(x, y, z)$ representa la distancia desde un punto (x, y, z) de C a un eje L, el momento de inercia I_L está definido por la integral de línea

$$I_L = \int_C \delta^2(x, y, z)\varphi(x, y, z)\, ds,$$

en donde $\varphi(x, y, z)$ es la densidad en (x, y, z). Los momentos de inercia respecto a los ejes coordenados se representan por I_x, I_y e I_z.

EJEMPLO 2. Calcular el momento de inercia I_z del muelle del ejemplo 1.

Solución. Aquí $\delta^2(x, y, z) = x^2 + y^2 = a^2$ y $\varphi(x, y, z) = x^2 + y^2 + z^2$, así que tenemos

$$I_z = \int_C (x^2 + y^2)(x^2 + y^2 + z^2)\, ds = a^2 \int_C (x^2 + y^2 + z^2)\, ds = Ma^2,$$

en donde M es la masa, como se calculó en el ejemplo 1.

10.9 Ejercicios

1. Un campo de fuerzas f del espacio de tres dimensiones viene dado por $f(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + (xz - y)\boldsymbol{k}$. Calcular el trabajo realizado por esa fuerza al mover una partícula desde $(0, 0, 0)$ a $(1, 2, 4)$ a lo largo del segmento de recta que une esos puntos.
2. Hallar el trabajo realizado por la fuerza $f(x, y) = (x^2 - y^2)\boldsymbol{i} + 2xy\boldsymbol{j}$ al mover una partícula en sentido contrario al de las agujas del reloj recorriendo una vez el contorno del cuadrado limitado por los ejes coordenados y las rectas $x = a$ e $y = a$, $a > 0$.
3. Un campo de fuerzas bidimensional f viene dado por la ecuación $f(x, y) = cxy\boldsymbol{i} + x^6y^2\boldsymbol{j}$ siendo c una constante positiva. Esa fuerza actúa sobre una partícula que se mueve desde $(0, 0)$ hasta la recta $x = 1$ siguiendo una curva de la forma

$$y = ax^b, \text{ en donde } a > 0 \quad \text{ y } \quad b > 0.$$

Encontrar el valor de a (en función de c) tal que el trabajo realizado por esa fuerza sea independiente de b.

4. Un campo de fuerzas f en el espacio de tres dimensiones viene dado por la fórmula $f(x, y, z) = yz\boldsymbol{i} + xz\boldsymbol{j} + x(y+1)\boldsymbol{k}$. Calcular el trabajo realizado por f al mover una partícula recorriendo una vez el contorno del triángulo de vértices $(0, 0, 0)$, $(1, 1, 1)$, $(-1, 1, -1)$ en este orden.
5. Calcular el trabajo realizado por el campo de fuerzas $f(x, y, z) = (y-z)\boldsymbol{i} + (z-x)\boldsymbol{j} + (x-y)\boldsymbol{k}$ a lo largo de la curva de intersección de la esfera $x^2 + y^2 + z^2 = 4$ y el plano $z = y \tan\theta$, en donde $0 < \theta < \pi/2$. El camino es recorrido de modo que, observado el plano xy desde el eje z positivo, el sentido aparezca contrario al de las agujas del reloj.
6. Calcular el trabajo realizado por el campo de fuerzas $f(x, y, z) = y^2\boldsymbol{i} + z^2\boldsymbol{j} + x^2\boldsymbol{k}$ a lo largo de la curva de intersección de la esfera $x^2 + y^2 + z^2 = a^2$ y el cilindro $x^2 + y^2 = ax$, siendo $z \geq 0$ y $a > 0$. El camino es recorrido de modo que, observado el plano xy desde el eje z positivo el sentido sea el de las agujas del reloj.

Calcular la integral de línea con respecto a la longitud de arco en cada uno de los ejercicios del 7 al 10.

7. $\int_C (x + y)\, ds$, siendo C el triángulo de vértices $(0, 0)$, $(1, 0)$ y $(0, 1)$, recorrido en sentido contrario al de las agujas del reloj.
8. $\int_C y^2 ds$, en donde C tiene la ecuación vectorial

$$\boldsymbol{\alpha}(t) = a(t - \operatorname{sen} t)\boldsymbol{i} + a(1 - \cos t)\boldsymbol{j}, \qquad 0 \leq t \leq 2\pi.$$

9. $\int_C (x^2 + y^2)\, ds$, donde C tiene la ecuación vectorial

$$\boldsymbol{\alpha}(t) = a(\cos t + t \operatorname{sen} t)\boldsymbol{i} + a(\operatorname{sen} t - t\cos t)\boldsymbol{j}, \qquad 0 \leq t \leq 2\pi.$$

10. $\int_C z\, ds$, donde C tiene la ecuación vectorial

$$\boldsymbol{\alpha}(t) = t\cos t\, \boldsymbol{i} + t \operatorname{sen} t\, \boldsymbol{j} + t\boldsymbol{k}, \qquad 0 \leq t \leq t_0.$$

11. Consideremos un alambre semicircular uniforme de radio a.
    a) Demostrar que el centroide está situado en el eje de simetría a una distancia $2a/\pi$ del centro.
    b) Demostrar que el momento de inercia respecto al diámetro que pasa por los extremos del alambre es $\frac{1}{2}Ma^2$, siendo M la masa del alambre.
12. Un alambre tiene la forma de un círculo $x^2 + y^2 = a^2$. Determinar su masa y su momento de inercia respecto a un diámetro si la densidad en (x, y) es $|x| + |y|$.
13. Hallar la masa de un alambre cuya forma es la de la curva de intersección de la esfera $x^2 + y^2 + z^2 = 1$ y el plano $x + y + z = 0$ si la densidad del alambre en (x, y, z) es x^2.
14. Un alambre uniforme tiene la forma de la porción de curva de intersección de las dos superficies $x^2 + y^2 = z^2$ e $y^2 = x$ que une los puntos $(0, 0, 0)$ y $(1, 1, \sqrt{2})$. Hallar la coordenada $\bar{z}$ del centroide.
15. Determinar las coordenadas $\bar{x}$ e $\bar{y}$ del centro de gravedad del muelle que se cita en el ejemplo 1 de la sección 10.8.
16. Para el muelle del ejercicio 1 de la sección 10.8, calcular los momentos de inercia I_x e I_y.

10.10 Conjuntos conexos abiertos. Independencia del camino

Sea S un conjunto abierto de $\mathbf{R}^n$. El conjunto S se llama *conexo* si todo par de puntos de S puede unirse mediante un camino regular a trozos cuya gráfica está situada en S. Esto es, para todo par de puntos $\boldsymbol{a}$ y $\boldsymbol{b}$ de S existe un camino regular a trozos $\boldsymbol{\alpha}$ definido en un intervalo $[a, b]$ tal que $\boldsymbol{\alpha}(t) \in S$ para cada t de $[a, b]$, siendo $\boldsymbol{\alpha}(a) = \boldsymbol{a}$ y $\boldsymbol{\alpha}(b) = \boldsymbol{b}$.

En la figura 10.3 se muestran tres ejemplos de conjuntos convexos abiertos. Ejemplos análogos a ésos en el espacio de tres dimensiones podrían ser a) un sólido elipsoidal, b) un sólido poliédrico, y c) un sólido tórico; en cada caso sólo se consideran los puntos interiores.

Un conjunto abierto S se dice que es *no conexo* si S es la reunión de dos o más conjuntos abiertos no vacíos disjuntos. En la figura 10.4 se muestra un ejemplo. Puede demostrarse que la clase de los conjuntos conexos abiertos es idéntica a la de los conjuntos abiertos que son no conexos. (*)

Sea ahora $\boldsymbol{f}$ un campo vectorial continuo en un conjunto conexo abierto S. Elijamos dos puntos $\boldsymbol{a}$ y $\boldsymbol{b}$ de S y consideremos la integral de línea de $\boldsymbol{f}$ a lo largo de un camino regular a trozos situado en S que una $\boldsymbol{a}$ y $\boldsymbol{b}$. El valor de la integral depende, en general, del camino que une $\boldsymbol{a}$ y $\boldsymbol{b}$. Para ciertos campos vectoriales, la integral depende únicamente de los extremos $\boldsymbol{a}$ y $\boldsymbol{b}$ y no del camino que los une. En este caso decimos que la integral es *independiente del camino que une* $\boldsymbol{a}$ *y* $\boldsymbol{b}$. Decimos que la integral de linea de $\boldsymbol{f}$ es *independiente del camino en* S si es independiente del camino que une $\boldsymbol{a}$ y $\boldsymbol{b}$ para todo par de puntos $\boldsymbol{a}$ y $\boldsymbol{b}$ de S.

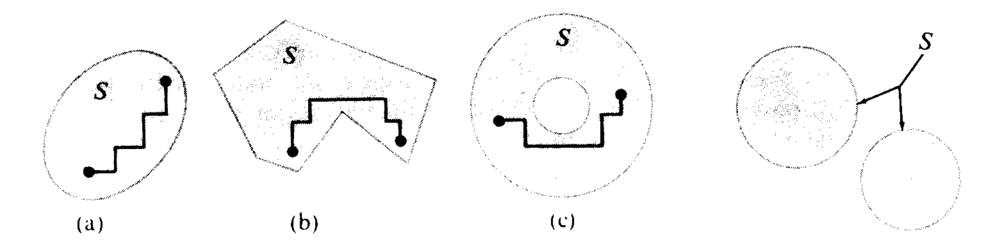


FIGURA 10.3 *Ejemplos de conjuntos conexos abiertos.*

FIGURA 10.4 *Conjunto no conexo S, reunión de dos discos circulares disjuntos.*

¿Qué campos vectoriales tienen integrales de línea independientes del camino? Para contestar esta pregunta, extendemos los teoremas fundamentales primero y segundo del cálculo a las integrales de línea.

(*) Para estudiar con mayor profundidad la conexión de conjuntos, véase el capítulo 8 de la obra del autor *Análisis Matemático,* Editorial Reverté, Barcelona.

10.11 Segundo teorema fundamental del cálculo para integrales de línea

El segundo teorema fundamental para funciones reales, como se demostró en el volumen I (teorema 5.3), establece que

$$\int_a^b \varphi'(t)\,dt = \varphi(b) - \varphi(a),$$

con tal que φ' sea continua en un cierto intervalo abierto que contenga a y b. Para extender ese resultado a las integrales de línea necesitamos una versión algo más fuerte del teorema en la que la continuidad de φ' se supone solamente en el intervalo abierto (a, b).

TEOREMA 10.2. *Si φ es una función real continua en un intervalo cerrado $[a, b]$, si suponemos que la integral $\int_a^b \varphi'(t)dt$ existe y si φ' es continua en el intervalo abierto (a, b), entonces tenemos*

$$\int_a^b \varphi'(t)\,dt = \varphi(b) - \varphi(a).$$

Demostración. Para cada x de $[a, b]$ definamos $f(x) = \int_a^x \varphi'(t)\,dt$. Queremos demostrar que

$$f(b) = \varphi(b) - \varphi(a). \tag{10.2}$$

Según el teorema 3.4 del volumen I, f es continua en el intervalo cerrado $[a, b]$. Según el teorema 5.1 del volumen I, f es derivable en el intervalo abierto (a, b), con $f'(x) = \varphi'(x)$ para cada x de (a, b). Por consiguiente, según el teorema de la derivada nula (teorema 5.2 del volumen I), la diferencia $f - \varphi$ es constante en el intervalo abierto (a, b). Por la continuidad, $f - \varphi$ también es constante en el intervalo cerrado $[a, b]$. En particular, $f(b) - \varphi(b) = f(a) - \varphi(a)$. Pero como $f(a) = 0$, esto demuestra (10.2).

TEOREMA 10.3. SEGUNDO TEOREMA FUNDAMENTAL DEL CÁLCULO PARA INTEGRALES DE LÍNEA. *Si φ es un campo escalar diferenciable con gradiente continuo $\nabla\varphi$ en un conjunto conexo abierto S en $\mathbf{R}^n$, entonces para dos puntos cualesquiera $\boldsymbol{a}$ y $\boldsymbol{b}$ unidos por un camino regular a trozos $\boldsymbol{\alpha}$ situado en S tenemos*

$$\int_{\boldsymbol{a}}^{\boldsymbol{b}} \nabla\varphi \cdot d\boldsymbol{\alpha} = \varphi(\boldsymbol{b}) - \varphi(\boldsymbol{a}).$$

Demostración. Elijamos dos puntos cualesquiera $\boldsymbol{a}$ y $\boldsymbol{b}$ de S y unámoslos con

un camino regular a trozos $\boldsymbol{\alpha}$ situado en S definido en un intervalo $[a, b]$. Supongamos primero que $\boldsymbol{\alpha}$ es *regular* en $[a, b]$. Entonces la integral de línea de $\nabla\varphi$ entre $\boldsymbol{a}$ y $\boldsymbol{b}$ a lo largo de $\boldsymbol{\alpha}$ viene dada por

$$\int_{\boldsymbol{a}}^{\boldsymbol{b}} \nabla\varphi \cdot d\boldsymbol{\alpha} = \int_a^b \nabla\varphi[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t)\, dt .$$

Según la regla de la cadena tenemos

$$\nabla\varphi[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t) = g'(t),$$

en donde g es la función compuesta definida en $[a, b]$ por la fórmula

$$g(t) = \varphi[\boldsymbol{\alpha}(t)] .$$

La derivada g' es continua en el intervalo abierto (a, b) debido a que $\nabla\varphi$ es continua en S y $\boldsymbol{\alpha}$ es regular. Por lo tanto podemos aplicar el teorema 10.2 a g obteniendo

$$\int_{\boldsymbol{a}}^{\boldsymbol{b}} \nabla\varphi \cdot d\boldsymbol{\alpha} = \int_a^b g'(t)\, dt = g(b) - g(a) = \varphi[\boldsymbol{\alpha}(b)] - \varphi[\boldsymbol{\alpha}(a)] = \varphi(\boldsymbol{b}) - \varphi(\boldsymbol{a}) .$$

Esto prueba el teorema si $\boldsymbol{\alpha}$ es regular.

Cuando $\boldsymbol{\alpha}$ es regular a trozos efectuamos una partición del intervalo $[a, b]$ en un número finito (por ejemplo r) de subintervalos $[t_{k-1}, t_k]$, en cada uno de los cuales $\boldsymbol{\alpha}$ es regular, y aplicamos el resultado que acabamos de demostrar a cada subintervalo. Esto nos dice

$$\int_{\boldsymbol{a}}^{\boldsymbol{b}} \nabla\varphi = \sum_{k=1}^{r} \int_{\boldsymbol{\alpha}(t_{k-1})}^{\boldsymbol{\alpha}(t_k)} \nabla\varphi = \sum_{k=1}^{r} \{\varphi[\boldsymbol{\alpha}(t_k)] - \varphi[\boldsymbol{\alpha}(t_{k-1})]\} = \varphi(\boldsymbol{b}) - \varphi(\boldsymbol{a}),$$

como queríamos demostrar.

Como consecuencia del teorema 10.3 vemos que la integral de línea de un gradiente es independiente del camino en cualquier conjunto S conexo en el que el gradiente sea continuo. Para un camino cerrado tenemos $\boldsymbol{b} = \boldsymbol{a}$, así que $\varphi(\boldsymbol{b}) - \varphi(\boldsymbol{a}) = 0$. Dicho de otro modo, *la integral de línea de un gradiente continuo es cero a lo largo de todo camino cerrado regular a trozos situado en S.* En la sección 10.14 demostraremos (en el teorema 10.4) que los gradientes son los *únicos* campos vectoriales continuos con esta propiedad.

10.12 Aplicaciones a la Mecánica

Si un campo vectorial f es el gradiente de un campo escalar φ, entonces φ se llama *función potencial* para f. En el espacio de tres dimensiones, los conjuntos de nivel φ se llaman *superficies equipotenciales;* en dos dimensiones se llaman *líneas equipotenciales.* (Si φ representa la temperatura, la palabra «equipotencial» se reemplaza por «isoterma»; si φ representa la presión se emplea la palabra «isobara».)

EJEMPLO 1. En el espacio tridimensional, pongamos $\varphi(x, y, z) = r^n$, siendo $r = (x^2 + y^2 + z^2)^{1/2}$. Para todo entero n tenemos

$$\nabla(r^n) = nr^{n-2}\boldsymbol{r},$$

donde $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$. (Véase el ejercicio 8 de la sección 8.14). Por lo tanto φ es un potencial del campo vectorial

$$\boldsymbol{f}(x, y, z) = nr^{n-2}\boldsymbol{r}.$$

Las superficies equipotenciales de φ son esferas concéntricas con el centro en el origen.

EJEMPLO 2. *Potencial newtoniano.* La ley de la gravitación de Newton establece que la fuerza $\boldsymbol{f}$ que ejerce una partícula de masa M sobre otra particula de masa m es un vector de longitud GmM/r^2, en donde G es una constante y r es la distancia entre las dos partículas. Situemos el origen en la partícula de masa M, y sea $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ el vector posición que une el origen a la partícula de masa m. Entonces $r = \|\boldsymbol{r}\|$ y $-\boldsymbol{r}/r$ es un vector unitario con la misma dirección que $\boldsymbol{f}$, con lo que la ley de Newton toma la forma

$$\boldsymbol{f} = -GmMr^{-3}\boldsymbol{r}.$$

Haciendo $n = -1$ en el ejemplo 1 vemos que la fuerza de gravitación $\boldsymbol{f}$ es el gradiente del campo escalar dado por

$$\varphi(x, y, z) = GmMr^{-1}.$$

Éste es el *potencial de Newton.*

El trabajo efectuado por la fuerza de gravitación al mover la partícula de masa m desde (x_1, y_1, z_1) a (x_2, y_2, z_2) es

$$\varphi(x_1, y_1, z_1) - \varphi(x_2, y_2, z_2) = GmM\left(\frac{1}{r_1} - \frac{1}{r_2}\right),$$

en donde $r_1 = (x_1^2 + y_1^2 + z_1^2)^{1/2}$ y $r_2 = (x_2^2 + y_2^2 + z_2^2)^{1/2}$. Si los dos puntos están en la misma superficie equipotencial entonces $r_1 = r_2$ y no se realiza trabajo alguno.

EJEMPLO 3. *Principio de conservación de la energía mecánica.* Sea $\boldsymbol{f}$ un campo de fuerzas continuo que tiene un potencial φ en un conjunto conexo abierto S. El teorema 10.3 nos dice que el trabajo efectuado por $\boldsymbol{f}$ al mover una partícula desde $\boldsymbol{a}$ hasta $\boldsymbol{x}$ siguiendo un camino regular a trozos situado en S es $\varphi(\boldsymbol{x}) - \varphi(\boldsymbol{a})$, la variación de la función potencial. En el ejemplo 2 de la sección 10.6 se demostró que este trabajo es también igual a la variación de la energía cinética de la partícula, $k(\boldsymbol{x}) - k(\boldsymbol{a})$ en donde $k(\boldsymbol{x})$ representa la energía cinética de la partícula cuando está situada en $\boldsymbol{x}$. Así pues, tenemos

$$k(\boldsymbol{x}) - k(\boldsymbol{a}) = \varphi(\boldsymbol{x}) - \varphi(\boldsymbol{a}),$$

o

$$k(\boldsymbol{x}) - \varphi(\boldsymbol{x}) = k(\boldsymbol{a}) - \varphi(\boldsymbol{a}). \tag{10.3}$$

El escalar $-\varphi(\boldsymbol{x})$ se denomina *energía potencial*(*) de la partícula.

Si $\boldsymbol{a}$ se mantiene fijo y $\boldsymbol{x}$ se hace variar en el conjunto S, la ecuación (10.3) nos dice que la suma de $k(\boldsymbol{x})$ y $-\varphi(\boldsymbol{x})$ es constante. Es decir, *si un campo de fuerzas es un gradiente, la suma de las energías cinética y potencial de una partícula que se desplaza en dicho campo es constante.* En Mecánica esto se llama principio de conservación de la energía (mecánica). Un campo de fuerzas con una función potencial se llama *conservativo* porque la energía total, cinética más potencial, se conserva. En un campo conservativo, no se realiza trabajo alguno al mover una partícula alrededor de una curva cerrada volviendo al punto de partida. Un campo de fuerzas no será conservativo si existe fricción o viscosidad en el sistema, puesto que ésas tienden a convertir energía mecánica en energía calorífica.

10.13 Ejercicios

1. Determinar cuáles de los siguientes conjuntos abiertos S de $\mathbf{R}^2$ son conexos. Para cada conjunto conexo, elegir dos puntos distintos cualesquiera de S y explicar cómo se podría encontrar en S una curva regular a trozos que los uniera.

   a) $S = \{(x, y) \mid x^2 + y^2 \geq 0\}$.
   b) $S = \{(x, y) \mid x^2 + y^2 > 0\}$.
   c) $S = \{(x, y) \mid x^2 + y^2 < 1\}$.
   d) $S = \{(x, y) \mid 1 < x^2 + y^2 < 2\}$.
   e) $S = \{(x, y) \mid x^2 + y^2 > 1 \quad \text{y} \quad (x - 3)^2 + y^2 > 1\}$.
   f) $S = \{(x, y) \mid x^2 + y^2 < 1 \quad \text{o} \quad (x - 3)^2 + y^2 < 1\}$.

(*) Algunos autores consideran a $-\varphi$ como la función potencial de $\boldsymbol{f}$ de modo que la energía potencial en $\boldsymbol{x}$ será igual al valor de la función potencial φ en $\boldsymbol{x}$.

2. Dado un campo vectorial bidimensional

$$f(x, y) = P(x, y)\boldsymbol{i} + Q(x, y)\boldsymbol{j},$$

en el que las derivadas parciales $\partial P/\partial y$ y $\partial Q/\partial x$ son continuas en un conjunto abierto S. Si f es el gradiente de un cierto potencial φ, demostrar que

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

en cada punto de S.

3. En cada uno de los siguientes campos vectoriales, aplicar el resultado del ejercicio 2 para demostrar que f *no es* un gradiente. Hallar seguidamente un camino cerrado C tal que $\oint_C f \neq 0$.
   a) $f(x, y) = y\boldsymbol{i} - x\boldsymbol{j}$.
   b) $f(x, y) = y\boldsymbol{i} + (xy - x)\boldsymbol{j}$.

4. Dado un campo vectorial tridimensional

$$f(x, y, z) = P(x, y, z)\boldsymbol{i} + Q(x, y, z)\boldsymbol{j} + R(x, y, z)\boldsymbol{k},$$

en el que las derivadas parciales

$$\frac{\partial P}{\partial y}, \frac{\partial P}{\partial z}, \frac{\partial Q}{\partial x}, \frac{\partial Q}{\partial z}, \frac{\partial R}{\partial x}, \frac{\partial R}{\partial y},$$

son continuas en un conjunto abierto S. Si f es el gradiente de una cierta función potencial φ, demostrar que

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \qquad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}$$

en cada punto de S.

5. En cada uno de los siguientes campos vectoriales, aplicar el resultado del ejercicio 4 para demostrar que f *no* es un gradiente. Hallar seguidamente un camino cerrado C tal que $\oint_C f \neq 0$.
   a) $f(x, y, z) = y\boldsymbol{i} + x\boldsymbol{j} + x\boldsymbol{k}$.
   b) $f(x, y, z) = xy\boldsymbol{i} + (x^2 + 1)\boldsymbol{j} + z^2\boldsymbol{k}$.

6. Un campo de fuerzas f está definido en el espacio de tres dimensiones por la ecuación

$$f(x, y, z) = y\boldsymbol{i} + z\boldsymbol{j} + yz\boldsymbol{k}.$$

   a) Determinar si f es o no conservativo.
   b) Calcular el trabajo realizado al mover una partícula a lo largo de la curva de ecuación

$$\boldsymbol{\alpha}(t) = \cos t\,\boldsymbol{i} + \operatorname{sen} t\,\boldsymbol{j} + e^t\boldsymbol{k}$$

cuando t varía de 0 a π.

7. Un campo de fuerzas bidimensional $\boldsymbol{F}$ tiene por ecuación

$$\boldsymbol{F}(x, y) = (x + y)\boldsymbol{i} + (x - y)\boldsymbol{j}.$$

a) Demostrar que el trabajo realizado por esa fuerza al mover una partícula siguiendo la curva

$$\boldsymbol{\alpha}(t) = f(t)\boldsymbol{i} + g(t)\boldsymbol{j}, \qquad a \leq t \leq b,$$

depende únicamente de $f(a)$, $f(b)$, $g(a)$, $g(b)$.
b) Hallar el trabajo realizado cuando $f(a) = 1$, $f(b) = 2$, $g(a) = 3$, $g(b) = 4$.

8. Un campo de fuerzas viene dado en coordenadas polares por la ecuación

$$\boldsymbol{F}(r,\theta) = -4 \operatorname{sen} \theta\, \boldsymbol{i} + 4 \operatorname{sen} \theta\, \boldsymbol{j}.$$

Calcular el trabajo efectuado al mover una partícula desde el punto $(1, 0)$ al origen siguiendo la espiral cuya ecuación polar es $r = e^{-\theta}$.

9. Un campo de fuerzas radial o «central» $\boldsymbol{F}$ en el plano puede expresarse en la forma $\boldsymbol{F}(x, y) = f(r)\boldsymbol{r}$ en donde $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$ y $r = \|\boldsymbol{r}\|$. Demostrar que un tal campo de fuerzas es conservativo.

10. Hallar el trabajo realizado por la fuerza $\boldsymbol{F}(x, y) = (3y^2 + 2)\boldsymbol{i} + 16x\boldsymbol{j}$ al mover una partícula desde $(-1, 0)$ a $(1, 0)$ siguiendo la mitad superior de la elipse $b^2x^2 + y^2 = b^2$. ¿Qué elipse (es decir, qué valor de b) hace mínimo el trabajo?

10.14 El primer teorema fundamental del cálculo para integrales de línea

En la sección 10.11 se extendió el segundo teorema fundamental del cálculo a las integrales de línea. En esta sección se extiende el primer teorema fundamental. Recordemos que éste establece que toda integral indefinida de una función continua f tiene una derivada igual a f. Esto es, si

$$\varphi(x) = \int_a^x f(t)\, dt,$$

en los puntos de continuidad de f tenemos

$$\varphi'(x) = f(x).$$

Para extender este teorema a las integrales de línea comenzamos con un campo vectorial $\boldsymbol{f}$, continuo en un conjunto conexo abierto S, y lo integramos a lo largo de una curva regular a trozos C entre un punto fijo $\boldsymbol{a}$ de S a un punto

cualquiera x. Designemos entonces con φ el campo escalar definido por la integral de línea

$$\varphi(\boldsymbol{x}) = \int_{\boldsymbol{a}}^{\boldsymbol{x}} \boldsymbol{f} \cdot d\boldsymbol{\alpha},$$

en donde $\boldsymbol{\alpha}$ es la función que origina C. Puesto que S es conexo, cada punto x de S puede ser alcanzado por una tal curva. Para que esta definición de $\varphi(x)$ carezca de ambigüedad, necesitamos saber que la integral depende únicamente de x y no del camino utilizado para unir $\boldsymbol{a}$ con x. Por consiguiente, es natural exigir que la integral de línea de f sea independiente del camino en S. En esas condiciones la extensión del primer teorema fundamental toma la forma siguiente:

TEOREMA 10.4. PRIMER TEOREMA FUNDAMENTAL PARA INTEGRALES DE LÍNEA. *Sea f un campo vectorial continuo en un conjunto conexo abierto S de $\mathbf{R}^n$, y supongamos que la integral de línea de f es independiente del camino en S. Sea $\boldsymbol{a}$ un punto fijo de S y definamos un campo escalar φ en S mediante la ecuación*

$$\varphi(\boldsymbol{x}) = \int_{\boldsymbol{a}}^{\boldsymbol{x}} \boldsymbol{f} \cdot d\boldsymbol{\alpha},$$

en donde $\boldsymbol{\alpha}$ es un camino regular a trozos de S que une $\boldsymbol{a}$ con x. Existe entonces el gradiente de φ y es igual a f; es decir

$$\nabla\varphi(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) \quad \text{para todo } \boldsymbol{x} \text{ de } S.$$

Demostración. Demostraremos que la derivada parcial $D_k\varphi(x)$ existe y es igual a $f_k(x)$, componente k-ésima de $f(x)$, para cada $k = 1, 2, \ldots, n$ y cada x de S.

Sea $B(x; r)$ una n-esfera con centro en x y radio r contenida en S. Si y es un vector unitario, el punto $x + hy$ también está contenido en S para todo h real que satisfaga la condición $0 < |h| < r$, y podemos formar el cociente de diferencias

$$\frac{\varphi(\boldsymbol{x} + h\boldsymbol{y}) - \varphi(\boldsymbol{x})}{h}.$$

En virtud de la propiedad aditiva de las integrales de línea, el numerador de ese cociente puede escribirse en la forma

$$\varphi(\boldsymbol{x} + h\boldsymbol{y}) - \varphi(\boldsymbol{x}) = \int_{\boldsymbol{x}}^{\boldsymbol{x}+h\boldsymbol{y}} \boldsymbol{f} \cdot d\boldsymbol{\alpha},$$

y el camino que une x con $x + hy$ puede ser cualquier camino regular a trozos

contenido en S. En particular, podemos emplear el segmento de recta representado por

$$\boldsymbol{\alpha}(t) = \boldsymbol{x} + th\boldsymbol{y}, \quad \text{en donde} \quad 0 \leq t \leq 1 .$$

Puesto que $\boldsymbol{\alpha}'(t) = h\boldsymbol{y}$ el cociente de diferencias toma la forma

$$\text{(10.4)} \qquad \frac{\varphi(\boldsymbol{x} + h\boldsymbol{y}) - \varphi(\boldsymbol{x})}{h} = \int_0^1 \boldsymbol{f}(\boldsymbol{x} + th\boldsymbol{y}) \cdot \boldsymbol{y}\, dt .$$

Tomemos ahora $\boldsymbol{y} = \boldsymbol{e}_k$, el k-ésimo vector coordenado unidad, y observemos que el integrando se transforma en $\boldsymbol{f}(\boldsymbol{x} + th\boldsymbol{y}) \cdot \boldsymbol{y} = f_k(\boldsymbol{x} + th\boldsymbol{e}_k)$. Hagamos luego el cambio de variable $u = ht$, $du = hdt$, y escribamos (10.4) en la forma

$$\text{(10.5)} \qquad \frac{\varphi(\boldsymbol{x} + h\boldsymbol{e}_k) - \varphi(\boldsymbol{x})}{h} = \frac{1}{h}\int_0^h f_k(\boldsymbol{x} + u\boldsymbol{e}_k)\, du = \frac{g(h) - g(0)}{h},$$

en donde g es la función definida en un intervalo abierto $(-r, r)$ por la ecuación

$$g(t) = \int_0^t f_k(\boldsymbol{x} + u\boldsymbol{e}_k)\, du .$$

Como cada componente f_k es continua en S, el primer teorema fundamental para integrales ordinarias nos dice que existe $g'(t)$ para cada t en $(-r, r)$ y que

$$g'(t) = f_k(\boldsymbol{x} + t\boldsymbol{e}_k) .$$

En particular, $g'(0) = f_k(\boldsymbol{x})$. Por consiguiente, si en (10.5) hacemos que $h \to 0$, encontramos que

$$\lim_{h \to 0} \frac{\varphi(\boldsymbol{x} + h\boldsymbol{e}_k) - \varphi(\boldsymbol{x})}{h} = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = g'(0) = f_k(\boldsymbol{x}) .$$

Esto demuestra que la derivada parcial $D_k\varphi(\boldsymbol{x})$ existe y es igual a $f_k(\boldsymbol{x})$, como queríamos demostrar.

10.15 Condiciones necesarias y suficientes para que un campo vectorial sea un gradiente

Los teoremas fundamentales primero y segundo para integrales de línea nos dicen que una condición necesaria y suficiente para que un campo vectorial sea

un gradiente en un conjunto conexo abierto es que su integral de línea entre dos puntos cualesquiera sea independiente del camino. Demostraremos ahora que esa condición equivale a afirmar que la integral de línea, a lo largo de cualquier camino *cerrado* regular a trozos, es nula. Todas esas condiciones están resumidas en el siguiente teorema.

TEOREMA 10.5. CONDICIONES NECESARIAS Y SUFICIENTES PARA QUE UN CAMPO VECTORIAL SEA UN GRADIENTE. *Si f es un campo vectorial continuo en un conjunto conexo abierto S de $\mathbf{R}^n$, entonces son equivalentes las tres afirmaciones siguientes:*

a) *f es gradiente de una cierta función potencial en S.*

b) *La integral de línea de f es independiente del camino en S.*

c) *La integral de línea de f alrededor de todo camino cerrado regular a trozos contenido en S es nula.*

Demostración. Demostraremos que b) implica a), a) implica c), y c) implica b). La afirmación b) implica a) en virtud del primer teorema fundamental. El segundo teorema fundamental pone de manifiesto que a) implica c).

Para completar la demostración probemos que c) implica b). Supongamos que se cumple c) y sean C_1 y C_2 dos curvas regulares a trozos cualesquiera en S con los mismos extremos. Sea la curva C_1 la gráfica de una función $\boldsymbol{\alpha}$ definida en un intervalo $[a, b]$, y C_2 la gráfica una función $\boldsymbol{\beta}$ definida en $[c, d]$.

Definamos una nueva función $\boldsymbol{\gamma}$ del modo siguiente:

$$\boldsymbol{\gamma}(t) = \begin{cases} \boldsymbol{\alpha}(t) & \text{si } a \le t \le b, \\ \boldsymbol{\beta}(b + d - t) & \text{si } b \le t \le b + d - c. \end{cases}$$

Entonces $\boldsymbol{\gamma}$ describe una curva cerrada C tal que

$$\oint_C f \cdot d\boldsymbol{\gamma} = \int_{C_1} f \cdot d\boldsymbol{\alpha} - \int_{C_2} f \cdot d\boldsymbol{\beta}.$$

Puesto que en virtud de c) $\oint_C f \cdot d\boldsymbol{\gamma} = 0$, tenemos $\int_{C_1} f \cdot d\boldsymbol{\alpha} = \int_{C_2} f \cdot d\boldsymbol{\beta}$, por lo que la integral de f es independiente del camino. Esto demuestra b). Así pues, a), b) y c) son equivalentes.

Observación: Si $\oint_C f \neq 0$ para una determinada curva cerrada C, entonces f *no es* un gradiente. No obstante, si una integral de línea $\oint_C f$ es cero para una cierta curva cerrada C o incluso para infinitas curvas cerradas, no se deduce necesariamente que f *es* un gradiente. Por ejemplo, el lector puede comprobar fácilmente que la integral de línea, del campo vectorial $f(x, y) = x\boldsymbol{i} + xy\boldsymbol{j}$ es cero para todo círculo con centro en el origen. No obstante, este campo vectorial no es un gradiente.

10.16 Condiciones necesarias para que un campo vectorial sea un gradiente

El primer teorema fundamental puede utilizarse para determinar si un campo vectorial dado es o no un gradiente en un conjunto conexo abierto S. Si la integral de línea de f es independiente del camino en S, definimos sencillamente un campo escalar φ integrando f desde un punto fijo a un punto x cualquiera de S siguiendo un camino adecuado en S. Calculamos entonces las derivadas parciales de φ y comparamos $D_k\varphi$ con f_k, k-ésimo componente de f. Si $D_k\varphi(x) = f_k(x)$ para todo x de S y todo k, entonces f es un gradiente en S y φ es un potencial. Si $D_k\varphi(x) \neq f_k(x)$ para algún k y un cierto x, entonces f *no* es un gradiente en S.

El teorema siguiente da otro criterio para determinar cuando un campo vectorial f *no* es un gradiente. Este criterio es muy útil en la práctica porque no exige integración alguna.

TEOREMA 10.6. CONDICIONES NECESARIAS PARA QUE UN CAMPO VECTORIAL SEA UN GRADIENTE. *Si $f = (f_1, \ldots, f_n)$ es un campo vectorial diferenciable con continuidad en un conjunto abierto S de $\mathbf{R}^n$, y si f es un gradiente en S, entonces las derivadas parciales de los componentes de f están ligados por las ecuaciones*

$$D_i f_j(x) = D_j f_i(x) \tag{10.6}$$

para $i, j = 1, 2, \ldots, n$ y todo x de S.

Demostración. Si f es un gradiente, $f = \nabla\varphi$ para una cierta función potencial φ. Esto significa que

$$f_j = D_j\varphi$$

para cada $j = 1, 2, \ldots, n$. Derivando ambos miembros de esta igualdad respecto a x_i encontramos

$$D_i f_j = D_i D_j \varphi .$$

Del mismo modo, tenemos

$$D_j f_i = D_j D_i \varphi .$$

Puesto que las derivadas parciales $D_i f_j$ y $D_j f_i$ son continuas en S, las dos derivadas parciales mixtas $D_i D_j \varphi$ y $D_j D_i \varphi$ deben ser iguales en S. Esto demuestra (10.6).

EJEMPLO 1. Determinar si el campo vectorial

$$f(x, y) = 3x^2 y \boldsymbol{i} + x^3 y \boldsymbol{j}$$

es o no un gradiente en algún subconjunto abierto de $\mathbf{R}^2$.

Solución. Tenemos aquí

$$f_1(x, y) = 3x^2y, \qquad f_2(x, y) = x^3y.$$

Las derivadas parciales D_2f_1 y D_1f_2 vienen dadas por

$$D_2f_1(x, y) = 3x^2, \qquad D_1f_2(x, y) = 3x^2y.$$

Puesto que $D_2f_1(x, y) \neq D_1f_2(x, y)$ excepto cuando $x = 0$ o $y = 1$, este campo vectorial no es un gradiente en ningún subconjunto abierto de $\mathbf{R}^2$.

El ejemplo que sigue pone de manifiesto que las condiciones del teorema 10.6 no siempre son suficientes para que un campo vectorial sea un gradiente.

EJEMPLO 2. Sea S el conjunto de todos los puntos $(x, y) \neq (0, 0)$ en $\mathbf{R}^2$, y sea $\boldsymbol{f}$ el campo vectorial definido en S por la ecuación

$$\boldsymbol{f}(x, y) = \frac{-y}{x^2 + y^2}\,\boldsymbol{i} + \frac{x}{x^2 + y^2}\,\boldsymbol{j}.$$

Demostrar que $D_1f_2 = D_2f_1$ en todo el conjunto S pero que, sin embargo, $\boldsymbol{f}$ no es un gradiente en S.

Solución. Es fácil comprobar que $D_1f_2(x, y) = D_2f_1(x, y)$ para todo punto (x, y) de S. (Véase ejercicio 17 de la sección 10.18.)

Para demostrar que $\boldsymbol{f}$ no es un gradiente en S calculamos la integral de línea de $\boldsymbol{f}$ a lo largo de la circunferencia unidad dada por

$$\boldsymbol{\alpha}(t) = \cos t\,\boldsymbol{i} + \operatorname{sen} t\,\boldsymbol{j}, \qquad 0 \leq t \leq 2\pi.$$

Obtenemos

$$\oint \boldsymbol{f} \cdot d\boldsymbol{\alpha} = \int_0^{2\pi} \boldsymbol{f}[\boldsymbol{\alpha}(t)] \cdot \boldsymbol{\alpha}'(t)\, dt = \int_0^{2\pi} (\operatorname{sen}^2 t + \cos^2 t)\, dt = 2\pi.$$

Puesto que la integral de línea a lo largo de ese camino cerrado no es cero, $\boldsymbol{f}$ no es un gradiente en S. En el ejercicio 18 de la sección 10.18 se discuten otras propiedades de ese campo vectorial.

Al final de este capítulo demostraremos que las condiciones necesarias del teorema 10.6 son también suficientes si se cumplen en un conjunto *convexo* abierto. (Véase el teorema 10.9.)

10.17 Métodos especiales para construir funciones potenciales

El primer teorema fundamental para las integrales de línea nos proporciona también un método para construir funciones potenciales. Si f es un gradiente continuo en un conjunto conexo abierto S, la integral de línea de f es independiente del camino en S. Por lo tanto podemos encontrar un potencial en S. Por lo tanto podemos encontrar un potencial φ integrando f desde un punto fijo $\boldsymbol{a}$ hasta un punto cualquiera $\boldsymbol{x}$ de S, siguiendo un camino regular a trozos situado en S. El campo escalar que así se obtiene depende de la elección del punto inicial $\boldsymbol{a}$. Si empezamos en otro punto inicial $\boldsymbol{b}$, obtenemos un nuevo potencial ψ. Pero, en virtud de la propiedad aditiva de las integrales de línea, φ y ψ pueden diferir únicamente en una constante, siendo ésta la integral de f entre $\boldsymbol{a}$ y $\boldsymbol{b}$.

Los ejemplos que siguen ilustran el empleo de distintos caminos de integración.

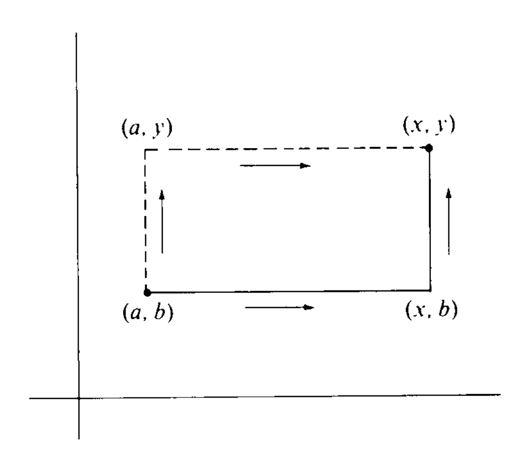


FIGURA 10.5 *Dos caminos poligonales que unen* (a, b) *con* (x, y).

EJEMPLO 1. *Construcción de un potencial en un rectángulo abierto.* Si f es un gradiente continuo en un rectángulo abierto de $\mathbf{R}^n$, puede construirse un potencial integrando entre un punto fijo y un punto arbitrario siguiendo un camino formado por segmentos rectilíneos paralelos a los ejes coordenados. En la figura 10.5, se representa un ejemplo en dos dimensiones. Podemos integrar primero desde (a, b) hasta (x, b) a lo largo de un segmento horizontal y luego desde (x, b) hasta (x, y) siguiendo un segmento vertical. A lo largo del segmento horizontal empleamos la representación paramétrica

$$\boldsymbol{\alpha}(t) = t\boldsymbol{i} + b\boldsymbol{j}, \qquad a \leq t \leq x,$$

y a lo largo del vertical empleamos esta otra

$$\boldsymbol{\alpha}(t) = x\boldsymbol{i} + t\boldsymbol{j}, \qquad b \leq t \leq y.$$

Si $f(x, y) = f_1(x, y)\boldsymbol{i} + f_2(x, y)\boldsymbol{j}$, la fórmula que resulta para un potencial $\varphi(x, y)$ es

$$\varphi(x, y) = \int_a^x f_1(t, b)\, dt + \int_b^y f_2(x, t)\, dt. \tag{10.7}$$

Podríamos asimismo integrar primero desde (a, b) hasta (a, y) siguiendo un segmento vertical y luego desde (a, y) hasta (x, y) a lo largo de un segmento horizontal como se indica con rectas punteadas en la figura 10.5. Esto nos da otra fórmula para $\varphi(x, y)$,

$$\varphi(x, y) = \int_b^y f_2(a, t)\, dt + \int_a^x f_1(t, y)\, dt. \tag{10.8}$$

Ambas fórmulas (10.7) y (10.8) nos dan el mismo valor para $\varphi(x, y)$ debido a que la integral de línea de un gradiente es independiente del camino.

EJEMPLO 2. *Construcción de una función potencial utilizando integrales indefinidas.* El uso de integrales indefinidas nos simplifican a menudo el cálculo de funciones potenciales. Por ejemplo, supongamos que un campo vectorial tridimensional $f = (f_1, f_2, f_3)$ es el gradiente de una función potencial φ en un conjunto abierto S de $\mathbf{R}^3$. Tenemos entonces

$$\frac{\partial \varphi}{\partial x} = f_1, \qquad \frac{\partial \varphi}{\partial y} = f_2, \qquad \frac{\partial \varphi}{\partial z} = f_3$$

en todo el conjunto S. Empleando integrales indefinidas e integrando la primera de esas ecuaciones respecto a x (manteniendo z e y constantes) encontramos

$$\varphi(x, y, z) = \int f_1(x, y, z)\, dx + A(y, z),$$

en donde $A(y, z)$ es una «constante de integración» que hay que determinar. Análogamente, si integramos la ecuación $\partial\varphi/\partial z = f_3$ respecto a y y a la ecuación $\partial\varphi/\partial y = f_2$ respecto a z obtenemos las relaciones

$$\varphi(x, y, z) = \int f_2(x, y, z)\, dy + B(x, z)$$

$$\varphi(x, y, z) = \int f_3(x, y, z)\, dz + C(x, y),$$

en las que $B(x, z)$ y $C(x, y)$ son funciones que tienen que determinarse. Encontrar φ significa encontrar tres funciones $A(y, z)$, $B(x, z)$, y $C(x, y)$ tales que las tres ecuaciones obtenidas para $\varphi(x, y, z)$ tengan sus segundos miembros coincidente. En muchos casos eso puede conseguirse a simple vista, como el siguiente ejemplo muestra.

EJEMPLO 3. Encontrar una función potencial φ para el campo vectorial definido en $\mathbf{R}^3$ por la ecuación

$$\boldsymbol{f}(x, y, z) = (2xyz + z^2 - 2y^2 + 1)\boldsymbol{i} + (x^2z - 4xy)\boldsymbol{j} + (x^2y + 2xz - 2)\boldsymbol{k}.$$

Solución. Sin conocer de antemano si $\boldsymbol{f}$ tiene o no función potencial φ, intentamos construir un potencial como se indicó en el ejercicio 2, suponiendo que tal potencial φ exista.

Integrando el componente f_1 respecto a x encontramos

$$\varphi(x, y, z) = \int (2xyz + z^2 - 2y^2 + 1)\, dx + A(y, z) = x^2yz + xz^2 - 2xy^2 + x + A(y, z).$$

Integrando f_2 respecto a y, y luego f_3 respecto a z, encontramos

$$\varphi(x, y, z) = \int (x^2z - 4xy)\, dy + B(x, z) = x^2yz - 2xy^2 + B(x, z),$$

$$\varphi(x, y, z) = \int (x^2y + 2xz - 2)\, dz + C(x, y) = x^2yz + xz^2 - 2z + C(x, y).$$

fácilmente se ve que eligiendo $A(y, z) = -2z$, $B(x, z) = xz^2 + x - 2z$ y $C(x, y) = x - 2xy^2$ coinciden las tres ecuaciones; luego la función φ dada por la ecuación

$$\varphi(x, y, z) = x^2yz + xz^2 - 2xy^2 + x - 2z$$

es un potencial para $\boldsymbol{f}$ en $\mathbf{R}^3$.

EJEMPLO 4. *Construcción de un potencial en un conjunto convexo.* Un conjunto S de $\mathbf{R}^n$ se llama *convexo* si todo par de puntos de S puede unirse con un segmento rectilíneo, cuyos puntos pertenecen todos a S. En la figura 10.6 se representa un ejemplo. Todo conjunto convexo abierto es conexo.

Si $\boldsymbol{f}$ es un gradiente en un conjunto convexo abierto, puede entonces construirse un potencial φ integrando $\boldsymbol{f}$ entre un punto fijo $\boldsymbol{a}$ de S hasta un punto

cualquiera x siguiendo un segmento de recta que los una. Tal segmento puede representarse paramétricamente por la función

$$\alpha(t) = a + t(x - a), \text{ en donde } 0 \leq t \leq 1.$$

Ésta nos da $\alpha'(t) = x - a$, así que el potencial correspondiente viene dado por la integral

(10.9) $$\varphi(x) = \int_0^1 f(a + t(x - a)) \cdot (x - a)\, dt.$$

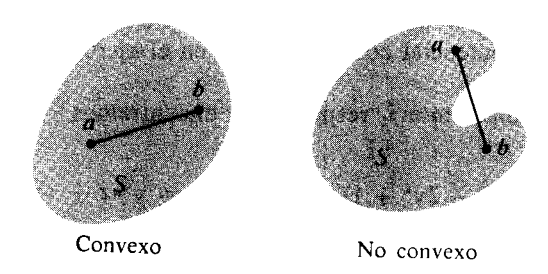


FIGURA 10.6 *En un conjunto convexo S, el segmento que une a con b está contenido en S cualesquiera que sean a y b de S.*

Si S contiene el origen podemos considerar $a = O$ y escribir (10.9) en la forma más sencilla

(10.10) $$\varphi(x) = \int_0^1 f(tx) \cdot x\, dt.$$

10.18 Ejercicios

En cada uno de los ejercicios del 1 al 12, está definido un campo vectorial f por las fórmulas que se dan. En cada caso determinar si f es o no el gradiente de un campo escalar. Cuando f sea un gradiente, hallar la correspondiente función potencial φ.

1. $f(x, y) = xi + yj$.
2. $f(x, y) = 3x^2yi + x^3j$.
3. $f(x, y) = (2xe^y + y)i + (x^2e^y + x - 2y)j$.
4. $f(x, y) = (\operatorname{sen} y - y \operatorname{sen} x + x)i + (\cos x + x \cos y + y)j$.
5. $f(x, y) = [\operatorname{sen}(xy) + xy \cos(xy)]i + x^2 \cos(xy)j$.
6. $f(x, y, z) = xi + yj + zk$.
7. $f(x, y, z) = (x + z)i - (y + z)j + (x - y)k$.

8. $f(x, y, z) = 2xy^3\boldsymbol{i} + x^2z^3\boldsymbol{j} + 3x^2yz^2\boldsymbol{k}$.
9. $f(x, y, z) = 3y^4z^2\boldsymbol{i} + 4x^3z^2\boldsymbol{j} - 3x^2y^2\boldsymbol{k}$.
10. $f(x, y, z) = (2x^2 + 8xy^2)\boldsymbol{i} + (3x^3y - 3xy)\boldsymbol{j} - (4y^2z^2 + 2x^3z)\boldsymbol{k}$.
11. $f(x, y, z) = (y^2 \cos x + z^3)\boldsymbol{i} - (4 - 2y \operatorname{sen} x)\boldsymbol{j} + (3xz^2 + 2)\boldsymbol{k}$.
12. $f(x, y, z) = (4xy - 3x^2z^2 + 1)\boldsymbol{i} + 2(x^2 + 1)\boldsymbol{j} - (2x^3z + 3z^2)\boldsymbol{k}$.

13. Un fluido se desplaza en el plano xy de modo que cada partícula se mueve en línea recta desde el origen. Si una partícula está a la distancia r del origen su velocidad es ar^n, en donde a y n son constantes.
a) Determinar los valores de a y n para los cuales el campo vectorial velocidad es el gradiente de un cierto campo escalar.
b) Encontrar una función potencial de la velocidad siempre que ésta sea un gradiente. El caso $n = -1$ debe tratarse separadamente.
14. Si φ y ψ son funciones potenciales para un campo vectorial continuo f en un conjunto conexo abierto S de $\mathbf{R}^n$, demostrar que $\varphi - \psi$ es constante en S.
15. Sea S el conjunto de todos los puntos $\boldsymbol{x} \neq \boldsymbol{O}$ de $\mathbf{R}^n$. Pongamos $r = \|\boldsymbol{x}\|$ y sea f el campo vectorial definido en S por la ecuación

$$f(\boldsymbol{x}) = r^p\boldsymbol{x},$$

siendo p una constante real. Hallar una función potencial para f en S. El caso $p = -2$ debe tratarse separadamente.
16. Resolver el ejercicio 15 para el campo vectorial definido por la ecuación

$$f(\boldsymbol{x}) = \frac{g'(r)}{r}\boldsymbol{x},$$

siendo g una función real con derivada continua en todo $\mathbf{R}^1$.

Los ejercicios que siguen se refieren al campo vectorial f definido en el conjunto S de todos los puntos $(x, y) \neq (0, 0)$ en $\mathbf{R}^2$ por la ecuación

$$f(x, y) = -\frac{y}{x^2 + y^2}\boldsymbol{i} + \frac{x}{x^2 + y^2}\boldsymbol{j}.$$

En el ejemplo 2 de la sección 10.6 se demostró que f no es un gradiente en S, aunque $D_1f_2 = D_2f_1$ en todo S.

17. Comprobar que para todo punto (x, y) de S tenemos

$$D_1f_2(x, y) = D_2f_1(x, y) = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

18. Este ejercicio muestra que f *es* un gradiente en el conjunto

$$T = \mathbf{R}^2 - \{(x, y) \mid y = 0, \quad x \leq 0\},$$

que consta de todos los puntos del plano xy excepto los del eje x no positivo.
a) Si $(x, y) \in T$, expresar x e y en coordenadas polares

$$x = r\cos\theta, \qquad y = r\,\text{sen}\,\theta,$$

en donde $r > 0$ y $-\pi < \theta < \pi$. Demostrar que θ viene dada por las fórmulas

$$\theta = \begin{cases} \arctan\dfrac{y}{x} & \text{si} \quad x > 0, \\ \dfrac{\pi}{2} & \text{si} \quad x = 0, \\ \arctan\dfrac{y}{x} + \pi & \text{si} \quad x < 0. \end{cases}$$

[Recuérdese la definición de la función arco tangente: Para cualquier valor real t, $\text{arc tan}\, t$ es el número real único que satisface las dos condiciones $\tan\varphi = t$ y $-\pi/2 < \varphi < \pi/2$.]
b) Deducir que

$$\frac{\partial\theta}{\partial x} = -\frac{y}{x^2 + y^2}, \qquad \frac{\partial\theta}{\partial y} = \frac{x}{x^2 + y^2}$$

para todo (x, y) en T. Esto demuestra que θ es una función potencial para f en el conjunto T.

Este ejercicio hace ver que la existencia de una función potencial depende no sólo del campo vectorial f sino también del conjunto en el que está definido el campo vectorial.

10.19 Aplicaciones a las ecuaciones diferenciales exactas de primer orden

Algunas ecuaciones diferenciales de primer orden pueden resolverse por medio de funciones potencial. Supongamos que tenemos una ecuación diferencial de la forma

$$y' = f(x, y).$$

Si multiplicamos ambos miembros por un factor no nulo $Q(x, y)$ transformamos esta ecuación en otra de la forma $Q(x, y)y' - f(x, y)Q(x,y) = 0$. Si ponemos $P(x, y)$ en lugar de $- f(x,y)Q(x, y)$ y empleamos la notación de Leibnitz para las derivadas, escribiendo y' en la forma dy/dx, la ecuación diferencial adopta la forma

$$P(x, y)\,dx + Q(x, y)\,dy = 0. \tag{10.11}$$

Supongamos que P y Q son continuas en un cierto conjunto abierto conexo S del plano. A una tal ecuación diferencial podemos asociar un campo vectorial $\boldsymbol{V}$, donde

$$\boldsymbol{V}(x, y) = P(x, y)\boldsymbol{i} + Q(x, y)\boldsymbol{j}.$$

Los componentes P y Q son los coeficientes de dx y dy en la ecuación (10.11). Se dice que la ecuación (10.11) es *exacta* en S si el campo vectorial $\boldsymbol{V}$ es el gradiente de un potencial; esto es, si $\boldsymbol{V}(x, y) = \nabla\varphi(x, y)$ para cada punto (x, y) de S; siendo φ un cierto campo escalar. Cuando un tal φ existe tenemos $\partial\varphi/\partial x = P$ y $\partial\varphi/\partial y = Q$, y la ecuación diferencial (10.11) se transforma en

$$\frac{\partial\varphi}{\partial x}\,dx + \frac{\partial\varphi}{\partial y}\,dy = 0.$$

Demostraremos ahora que cada solución de esta ecuación diferencial satisface la relación $\varphi(x, y) = C$, siendo C una constante. Precisando más, supongamos que existe una solución Y de la ecuación diferencial (10.11) definida en un intervalo abierto (a, b) tal que el punto $(x, Y(x))$ está en S para cada x de (a, b). Demostraremos que

$$\varphi[x, Y(x)] = C$$

para una cierta constante C. A tal fin introducimos la función compuesta g definida en (a, b) por la ecuación

$$g(x) = \varphi[x, Y(x)].$$

Según la regla de la cadena, la derivada de g viene dada por

$$(10.12)\qquad g'(x) = D_1\varphi[x, Y(x)] + D_2\varphi[x, Y(x)]\,Y'(x) = P(x, y) + Q(x, y)y',$$

donde $y = Y(x)$ e $y' = Y'(x)$. Si y satisface (10.11), $P(x, y) + Q(x, y)y' = 0$, de modo que $g'(x) = 0$ para cada x en (a, b) y, por consiguiente, g es constante en (a, b). Esto prueba que toda solución y satisface la ecuación $\varphi(x, y) = C$.

Podemos ahora invertir el razonamiento para hallar una solución de la ecuación diferencial. Sea la ecuación

$$(10.13)\qquad \varphi(x, y) = C$$

consideremos y como función diferenciable de x, tal como $y = Y(x)$ para valores de x de un intervalo (a, b), y sea $g(x) = \varphi[x, Y(x)]$. La ecuación (10.13) implica que g

es constante en (a, b). Luego, según (10.12), $P(x, y) + Q(x, y)y' = 0$, de modo que y es una solución. Por consiguiente, hemos demostrado el teorema siguiente:

TEOREMA 10.7. *Supongamos que la ecuación diferencial*

$$P(x, y)\,dx + Q(x, y)\,dy = 0 \tag{10.14}$$

es exacta en un conjunto conexo abierto S, y sea φ un campo escalar que satisface

$$\frac{\partial \varphi}{\partial x} = P \qquad \text{y} \qquad \frac{\partial \varphi}{\partial y} = Q$$

en todo S. Entonces toda solución $y = Y(x)$ de (10.14) *cuya gráfica está situada en S satisface la ecuación $\varphi[x, Y(x)] = C$ para un cierto valor de C. Recíprocamente, si la ecuación*

$$\varphi(x, y) = C$$

define y como función implícita diferenciable de x, entonces esta función es una solución de la ecuación diferencial (10.14).

El teorema precedente nos proporciona un método rápido para resolver ecuaciones diferenciales exactas de primer orden. Construimos una función potencial φ y escribimos la ecuación $\varphi(x, y) = C$, siendo C una constante. Siempre que esta ecuación defina y como función implícita de x, la correspondiente y satisface (10.14). Por tanto, podemos considerar la ecuación (10.13) como la representación de una familia de curvas integrales simplemente infinita, o sea, dependiente de un solo parámetro. Naturalmente, que los únicos valores admisibles de C son aquellos para los que $\varphi(x_0, y_0) = C$ para algún (x_0, y_0) de S.

EJEMPLO 1. Consideremos la ecuación diferencial

$$\frac{dy}{dx} = -\frac{3x^2 + 6xy^2}{6x^2y + 4y^3}.$$

Quitando denominadores se obtiene:

$$(3x^2 + 6xy^2)\,dx + (6x^2y + 4y^3)\,dy = 0.$$

Este es ahora un caso particular de (10.14) siendo $P(x, y) = 3x^2 + 6xy^2$ y $Q(x, y) = 6x^2y + 4y^3$. Integrando P con respecto a x y Q con respecto a y y

comparando los resultados, encontramos que una función potencial φ viene dada por la fórmula

$$\varphi(x, y) = x^3 + 3x^2y^2 + y^4 .$$

Según el teorema 10.7, cada solución diferencial satisface

$$x^3 + 3x^2y^2 + y^4 = C$$

para un cierto C. Esto representa en forma implícita una familia de curvas integrales. En este caso particular la ecuación es cuadrática en y^2 y puede despejarse y en función de x y de C.

EJEMPLO 2. Sea la ecuación diferencial de primer orden

$$y\,dx + 2x\,dy = 0 . \tag{10.15}$$

En este ejemplo $P(x, y) = y$ y $Q(x, y) = 2x$. Puesto que $\partial P/\partial y = 1$ y $\partial Q/\partial x = 2$, esta ecuación diferencial no es exacta. Sin embargo, si multiplicamos ambos miembros por y obtenemos una ecuación que *es* exacta:

$$y^2\,dx + 2xy\,dy = 0 . \tag{10.16}$$

Un potencial del campo vectorial $y^2\boldsymbol{i} + 2xy\boldsymbol{j}$ es $\varphi(x, y) = xy^2$, y toda solución de (10.16) satisface la relación $xy^2 = C$. Esta relación también representa una familia de curvas integrales de la ecuación (10.15).

El multiplicador y que ha convertido (10.15) en una ecuación exacta se llama *factor integrante*. En general, si al multiplicar una ecuación lineal de primer orden por un factor no nulo $\mu(x, y)$ se obtiene una ecuación diferencial exacta, el multiplicador $\mu(x, y)$ se denomina factor integrante de la ecuación original. Una ecuación diferencial puede tener más de un factor integrante. Por ejemplo, $\mu(x, y) = 2xy^3$ es otro factor integrante de (10.15). En los ejercicios que siguen se discuten algunos casos particulares de ecuaciones diferenciales para las que se pueden encontrar fácilmente factores integrantes.

10.20 Ejercicios

Demostrar que las ecuaciones diferenciales de los ejercicios del 1 al 5 son exactas, y en cada caso hallar una familia simplemente infinita de curvas integrales.

1. $(x + 2y)\,dx + (2x + y)\,dy = 0$.
2. $2xy\,dx + x^2\,dy = 0$.
3. $(x^2 - y)\,dx - (x + \operatorname{sen}^2 y)\,dy = 0$.

4. $4 \operatorname{sen} x \operatorname{sen} 3y \cos x \, dx - 3 \cos 3y \cos 2x \, dy = 0$.
5. $(3x^2y + 8xy^2)\, dx + (x^3 + 8x^2y + 12ye^y)\, dy = 0$.
6. Demostrar que una ecuación lineal de primer orden $y' + P(x)y = Q(x)$, posee el factor integrante $\mu(x) = e^{\int P(x)\,dx}$. Resolver con él la ecuación.
7. Sea $\mu(x, y)$ un factor integrante de la ecuación diferencial $P(x, y)\, dx + Q(x, y)\, dy = 0$. Demostrar que

$$\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} = Q \frac{\partial}{\partial x} \log |\mu| - P \frac{\partial}{\partial y} \log |\mu| .$$

Deducir de esta ecuación las siguientes reglas para encontrar factores integrantes:
a) Si $(\partial P/\partial y - \partial Q/\partial x)/Q$ es un función $f(x)$ de la sola variable x, $e^{\int f(x)\,dx}$ es un factor integrante.
b) Si $(\partial Q/\partial x - \partial P/\partial y)/P$ es una función $g(y)$ de la sola variable y, $e^{\int g(y)\,dy}$ es un factor integrante.

8. Como aplicación del ejercicio 7 hallar factores integrantes, y hallar las familias de curvas integrales para las ecuaciones siguientes:
a) $y\, dx - (2x + y)\, dy = 0$.
b) $(x^3 + y^3)\, dx - xy^2\, dy = 0$.
9. Si $\partial P/\partial y - \partial Q/\partial x = f(x)Q(x, y) - g(y)P(x, y)$, demostrar que $\{\int f(x)\, dx + \int g(y)\, dy\}$ es un factor integrante de la ecuación diferencial $P(x, y)\, dx + Q(x, y)\, dy = 0$. Hallar el correspondiente factor integrante para cada una de las ecuaciones siguientes y su familia simplemente infinita de curvas integrales:
a) $(2x^2y + y^2)\, dx + (2x^3 - xy)\, dy = 0$.
b) $(e^x \sec y - \tan y)\, dx + dy = 0$.
10. Las siguientes ecuaciones diferenciales tienen un factor integrante común. Hallarlo y determinar la familia de curvas integrales para cada ecuación.

$$(3y + 4xy^2)\, dx + (4x + 5x^2y)\, dy = 0,$$

$$(6y + x^2y^2)\, dx + (8x + x^3y)\, dy = 0.$$

10.21 Funciones de potencial en conjuntos convexos

En el teorema 10.6 se demostró que las condiciones

$$D_i f_j(x) = D_j f_i(x)$$

son *necesarias* para que un campo vectorial diferenciable con continuidad $f = (f_1, \ldots, f_n)$ sea un gradiente en un conjunto S de $\mathbf{R}^n$. Se demostró luego, con un ejemplo, que esas condiciones no eran siempre suficientes. Demostramos en esta sección que las condiciones son suficientes siempre que S sea convexo. La demostración hará uso del siguiente teorema relativo a la derivación bajo el signo integral.

TEOREMA 10.8. *Sea S un intervalo cerrado en $\mathbf{R}^n$ con interior no vacío y sea $J = [a, b]$ un intervalo cerrado en $\mathbf{R}^1$. Sea J_{n+1} el intervalo cerrado $S \times J$ de $\mathbf{R}^{n+1}$. Cada punto de J_{n+1} lo representamos por (x, t), en donde $x \in S$ y $t \in J$,*

$$(x, t) = (x_1, \ldots, x_n, t).$$

Supongamos que ψ es un campo escalar definido en J_{n+1} tal que la derivada parcial $D_k\psi$ sea continua en J_{n+1}, siendo k uno de los números $1, 2, \ldots, n$. Definamos un campo escalar φ en S por la ecuación

$$\varphi(x) = \int_a^b \psi(x, t)\, dt .$$

Existe entonces la derivada parcial $D_k\varphi$ en cada punto interior de S y se obtiene con la fórmula

$$D_k\varphi(x) = \int_a^b D_k\psi(x, t)\, dt .$$

Es decir, tenemos

$$D_k \int_a^b \psi(x, t)\, dt = \int_a^b D_k\psi(x, t)\, dt .$$

Observación. Este teorema nos indica que podemos derivar bajo el signo integral.

Demostración. Elijamos un x cualquiera en el interior de S. Puesto que int S es abierto, existe un $r > 0$ tal que $x + he_k \in$ int S para todo h que satisfaga $0 < h < r$ El vector e_k es el k-ésimo vector coordenado unidad de $\mathbf{R}^n$. Para tal h tenemos

$$(10.17) \qquad \varphi(x + he_k) - \varphi(x) = \int_a^b \{\psi(x + he_k, t) - \psi(x, t)\}\, dt .$$

Aplicando el teorema del valor medio al integrando tenemos

$$\psi(x + he_k, t) - \psi(x, t) = h\, D_k\psi(z, t),$$

en donde z pertenece al segmento rectilíneo que une x con $x + he_k$. Luego (10.17) se transforma en

$$\frac{\varphi(x + he_k) - \varphi(x)}{h} = \int_a^b D_k\psi(z, t)\, dt .$$

Por consiguiente

$$\frac{\varphi(x + he_k) - \varphi(x)}{h} - \int_a^b D_k\psi(x, t)\,dt = \int_a^b \{D_k\psi(z, t) - D_k\psi(x, t)\}\,dt.$$

La última integral tiene un valor absoluto que no excede

$$\int_a^b |D_k\psi(z, t) - D_k\psi(x, t)|\,dt \leq (b - a)\max|D_k\psi(z, t) - D_k\psi(x, t)|$$

en donde el máximo es alcanzado para todo z del segmento que une x con $x + he_k$, y todo t de $[a, b]$. Teniendo en cuenta la continuidad uniforme de D_k en $S \times J$ (teorema 9.10) llegamos a la conclusión de que para todo $\epsilon > 0$ existe un $\delta > 0$ tal que ese máximo es $< \epsilon/(b - a)$, siempre que $0 < |h| < \delta$. Por lo tanto

$$\left|\frac{\varphi(x + he_k) - \varphi(x)}{h} - \int_a^b D_k\psi(x, t)\,dt\right| < \epsilon \quad \text{siempre que} \quad 0 < |h| < \delta.$$

Esto demuestra que $D_k\varphi(x)$ existe y es igual a $\int_a^b D_k\psi(x, t)\,dt$, como queríamos demostrar.

Utilizaremos ahora este teorema para dar la siguiente condición necesaria y suficiente para que un campo vectorial sea un gradiente en un conjunto convexo.

TEOREMA 10.9. Sea $f = (f_1, \ldots, f_n)$ *un campo vectorial diferenciable con continuidad en un conjunto convexo abierto* S *de* $\mathbf{R}^n$. *El campo* f *es un gradiente en* S *si y sólo si*

(10.18) $$D_k f_j(x) = D_j f_k(x)$$

para cada x *de* S *y todos los índices* $k, j = 1, 2, \ldots, n$.

Demostración. Sabemos, por el teorema 10.6, que las condiciones son necesarias. Para demostrar la suficiencia, supongamos (10.18) y construyamos un potencial φ en S.

Supongamos, para simplificar, que S contenga el origen. Sea $\varphi(x)$ la integral de f a lo largo del segmento de recta que une O con un punto cualquiera x de S. Como antes se ha visto en la ecuación (10.10) tenemos

$$\varphi(x) = \int_0^1 f(tx)\cdot x\,dt = \int_0^1 \psi(x, t)\,dt,$$

en donde $\psi(\boldsymbol{x}, t) = \boldsymbol{f}(t\boldsymbol{x}) \cdot \boldsymbol{x}$. Existe un subintervalo cerrado n-dimensional T de S con interior no vacío tal que ψ satisface las hipótesis del teorema 10.8 en $T \times J$, siendo $J = [0, 1]$. Por consiguiente la derivada parcial $D_k\varphi(\boldsymbol{x})$ existe para cada valor $k = 1, 2, \ldots, n$ y puede calcularse derivando bajo el signo integral,

$$D_k\varphi(\boldsymbol{x}) = \int_0^1 D_k\psi(\boldsymbol{x}, t)\, dt.$$

Para calcular $D_k\psi(\boldsymbol{x}, t)$, derivamos el producto escalar $\boldsymbol{f}(t\boldsymbol{x}) \cdot \boldsymbol{x}$ y obtenemos

$$\begin{aligned} D_k\psi(\boldsymbol{x}, t) &= \boldsymbol{f}(t\boldsymbol{x}) \cdot D_k\boldsymbol{x} + D_k\{\boldsymbol{f}(t\boldsymbol{x})\} \cdot \boldsymbol{x} \\ &= \boldsymbol{f}(t\boldsymbol{x}) \cdot \boldsymbol{e}_k + t(D_kf_1(t\boldsymbol{x}), \ldots, D_kf_n(t\boldsymbol{x})) \cdot \boldsymbol{x} \\ &= f_k(t\boldsymbol{x}) + t(D_1f_k(t\boldsymbol{x}), \ldots, D_nf_k(t\boldsymbol{x})) \cdot \boldsymbol{x}, \end{aligned}$$

habiendo utilizado en el último paso la relación (10.18). Tenemos por lo tanto

$$D_k\psi(\boldsymbol{x}, t) = f_k(t\boldsymbol{x}) + t\nabla f_k(t\boldsymbol{x}) \cdot \boldsymbol{x}.$$

Pongamos ahora $g(t) = f_k(t\boldsymbol{x})$. Según la regla de la cadena tenemos

$$g'(t) = \nabla f_k(t\boldsymbol{x}) \cdot \boldsymbol{x}$$

así que la última fórmula que da $D_k\psi(\boldsymbol{x}, t)$ toma la forma

$$D_k\psi(\boldsymbol{x}, t) = g(t) + tg'(t).$$

Integrándola entre 0 y 1 encontramos

$$(10.19) \qquad D_k\varphi(\boldsymbol{x}) = \int_0^1 D_k\psi(\boldsymbol{x}, t) = \int_0^1 g(t)\, dt + \int_0^1 tg'(t)\, dt.$$

Integremos por partes la última integral y obtenemos

$$\int_0^1 tg'(t)\, dt = tg(t)\Big|_0^1 - \int_0^1 g(t)\, dt = g(1) - \int_0^1 g(t)\, dt.$$

Por consiguiente (10.19) toma la forma

$$D_k\varphi(\boldsymbol{x}) = g(1) = f_k(\boldsymbol{x}).$$

Esto demuestra que $\nabla\varphi = \boldsymbol{f}$ en S, lo cual completa la demostración.

11

INTEGRALES MÚLTIPLES

11.1 Introducción

En el volumen I se estudiaron las integrales $\int_a^b f(x)\,dx$, primero para funciones definidas y acotadas en intervalos finitos, y luego para funciones no acotadas e intervalos infinitos. En el capítulo 10 del volumen II se ha generalizado el concepto de integral introduciendo las integrales de línea. Este capítulo extiende el concepto en otra dirección. El intervalo unidimensional $[a, b]$ se reemplaza por un conjunto Q bi-dimensional, llamado *región de integración.* Consideremos primero regiones rectangulares; luego consideraremos regiones más generales con contornos curvilíneos. El integrando es un campo escalar f definido y acotado en Q. La integral que resulta se llama *integral doble* y se representa mediante los símbolos

$$\iint_Q f, \qquad \text{o} \qquad \iint_Q f(x, y)\,dx\,dy.$$

Como en el caso uni-dimensional, los símbolos dx y dy no desempeñan ningún papel en la definición de la integral doble; no obstante, son útiles en el cálculo y transformación de integrales.

El programa de este capítulo consta de varias partes. Discutimos primero la definición de integral doble. La introducción es análoga al caso uni-dimensional tratado en el volumen I. Se define la integral primero para funciones escalonadas y luego para funciones más generales. Como allí, la definición no proporciona un método para el cálculo efectivo de las integrales. Veremos que gran número de integrales dobles que se presentan en la práctica pueden calcularse por integración unidimensional reiterada. Encontraremos también una conexión entre las integrales dobles y las de línea. Asimismo se dan aplicaciones de las integrales dobles a problemas de áreas, volúmenes, masas, centros de gravedad y conceptos

relacionados con aquéllos. Por último indicamos el modo de extender los conceptos al espacio de n dimensiones.

11.2 Particiones de rectángulos. Funciones escalonadas

Sea Q un rectángulo, producto cartesiano de dos intervalos cerrados $[a, b]$ y $[c, d]$,

$$Q = [a, b] \times [c, d] = \{(x, y) \mid x \in [a, b] \quad \text{e} \quad y \in [c, d]\}.$$

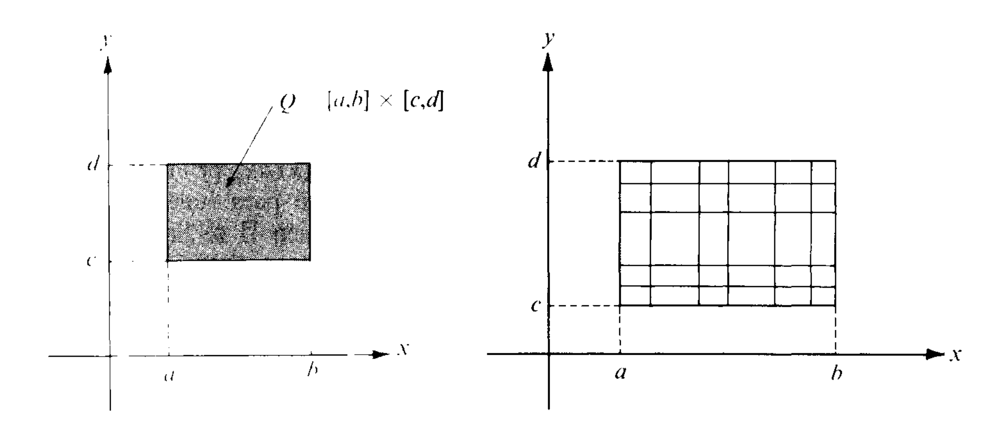


FIGURA 11.1 *Rectángulo Q producto cartesiano de dos intervalos.*

FIGURA 11.2 *Partición de un rectángulo Q.*

La figura 11.1 muestra un ejemplo. Consideremos dos particiones P_1 y P_2 de $[a, b]$ y $[c, d]$, respectivamente,

$$P_1 = \{x_0, x_1, \ldots, x_{n-1}, x_n\} \qquad \text{y} \qquad P_2 = \{y_0, y_1, \ldots, y_{m-1}, y_m\},$$

donde $x_0 = a$, $x_n = b$, $y_0 = c$, $y_m = d$. Se dice que el producto cartesiano $P_1 \times P_2$ es una partición de Q. Puesto que P_1 descompone $[a, b]$ en n subintervalos y P_2 descompone $[c, d]$ en m subintervalos, la partición $P = P_1 \times P_2$ descompone Q en mn subrectángulos. La figura 11.2 muestra un ejemplo de partición de Q en 30 subrectángulos. Una partición P' de Q se llama más fina que P si $P \subseteq P'$, esto es, si todo punto de P pertenece también a P'.

El producto cartesiano de dos subintervalos abiertos de P_1 y P_2 es un subrectángulo abierto (sin lados). Se llama subrectángulo abierto de P o de Q.

DEFINICIÓN DE FUNCIÓN ESCALONADA. *Una función f definida en un rectángulo Q se llama escalonada si existe una partición P de Q tal que f es constante en cada uno de los subrectángulos abiertos de P.*

En la figura 11.3 se muestra un ejemplo de una tal función escalonada. Una función escalonada también posee valores precisos en cada uno de los puntos

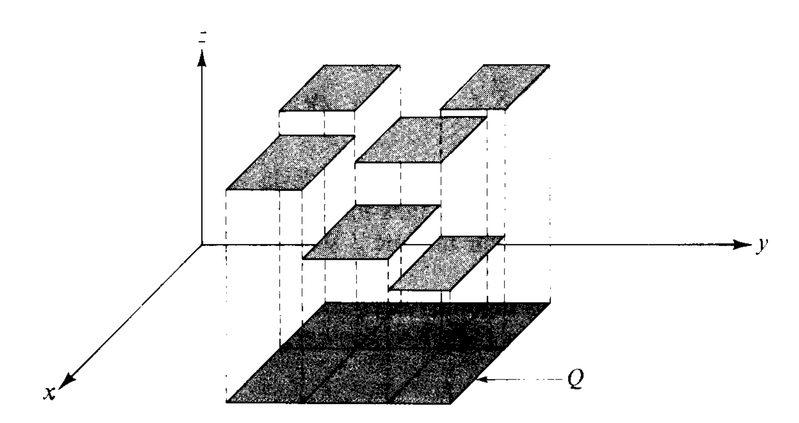


FIGURA 11.3 *Gráfica de una función escalonada definida en un rectángulo Q.*

de los contornos de los subrectángulos, pero los valores en esos puntos no tienen importancia en la teoría de la integración.

Si f y g son dos funciones escalonadas definidas en un rectángulo dado Q, la combinación lineal $c_1f + c_2g$ también es una función escalonada. En efecto, si P y P son particiones de Q tales que f es constante en los subrectángulos abiertos de P y g es constante en los subrectángulos abiertos de P', entonces $c_1f + c_2g$ es constante en los subrectángulos de la reunión $P \cup P'$ (que podemos llamar afinamiento de P y P'). Así pues, el conjunto de funciones escalonadas definidas en Q forma un espacio lineal.

11.3 Integral doble de una función escalonada

Sean $P = P_1 \times P_2$ una partición del rectángulo Q en mn subrectángulos y f una función escalonada, esto es, constante en los subrectángulos abiertos de Q. Designemos por Q_{ij} el subrectángulo determinado por $[x_{i-1}, x_i]$ e $[y_{j-1}, y_j]$ y sea c_{ij} el valor constante que toma f en los puntos interiores de Q_{ij}. Si f es positiva, el volumen de la caja rectangular con base Q_{ij} y altura c_{ij} es el producto

$$c_{ij} \cdot (x_i - x_{i-1})(y_j - y_{j-1}) .$$

para toda función escalonada f, positiva o no, la suma de todos esos productos se toma como definición de integral doble de f sobre Q. Así pues, tenemos la siguiente definición.

DEFINICIÓN DE LA INTEGRAL DOBLE DE UNA FUNCIÓN ESCALONADA. *Sea f una función escalonada que toma el valor constante c_{ij} en el subrectángulo abierto $(x_{i-1}, x_i) \times (y_{j-1}, y_j)$ de un rectángulo Q. La integral doble de f en Q está definida por la fórmula*

$$\iint_Q f = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \cdot (x_i - x_{i-1})(y_j - y_{j-1}). \tag{11.1}$$

Como en el caso unidimensional, el valor de la integral no varía si la partición P se sustituye por otra partición más fina P'. Así, pues, dicho valor es independiente de la elección de P con tal que f sea constante en los subrectángulos abiertos de Q.

Por brevedad, escribimos algunas veces Δx_i en lugar de $(x_i - x_{i-1})$ y Δy_j en lugar de $(y_j - y_{j-1})$, y la suma (11.1) se convierte en

$$\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \, \Delta x_i \, \Delta y_j .$$

Para recordar cómo se forma esta suma, podemos escribir la integral en la forma

$$\iint_Q f(x, y) \, dx \, dy .$$

Este símbolo no es más que otra notación para $\iint_Q f$.

Obsérvese que si f es constante en el interior de Q, es decir $f(x, y) = k$ si $a < x < b$ y $c < y < d$, tenemos

$$\iint_Q f = k(b - a)(d - c), \tag{11.2}$$

prescindiendo de los valores de f sobre los lados de Q. Puesto que

$$b - a = \int_a^b dx \qquad \text{y} \qquad d - c = \int_c^d dy ,$$

la fórmula (11.2) puede también escribirse así

$$\iint_Q f = \int_c^d \left[\int_a^b f(x, y) \, dx \right] dy = \int_a^b \left[\int_c^d f(x, y) \, dy \right] dx . \tag{11.3}$$

Las integrales que aparecen en el segundo miembro son unidimensionales, y se dice que la fórmula proporciona el valor de la integral doble por integración *repetida* o por integraciones *sucesivas*. En particular, si aplicamos esta fórmula cuando f es una función escalonada del tipo antes descrito, podemos escribir:

$$\iint_{Q_{ij}} f = \int_{y_{j-1}}^{y_j} \left[\int_{x_{i-1}}^{x_i} f(x, y)\, dx \right] dy = \int_{x_{i-1}}^{x_i} \left[\int_{y_{j-1}}^{y_j} f(x, y)\, dy \right] dx .$$

Sumando respecto a i y j y usando (11.1), encontramos que (11.3) es válida para funciones escalonadas.

Las otras propiedades que siguen de la integral doble de una función escalonada son generalizaciones de los correspondientes teoremas unidimensionales. Pueden demostrarse como consecuencias inmediatas de la definición (11.1) o mediante la fórmula (11.3) y los teoremas correspondientes para las integrales unidimensionales. En los teoremas siguientes los símbolos s y t representan funciones escalonadas definidas en un rectángulo Q. Para eludir casos particulares triviales suponemos que Q es un rectángulo no degenerado; en otras palabras, que Q no es tan sólo un punto o un segmento rectilíneo.

TEOREMA 11.1. LINEALIDAD. *Cualesquiera que sean los números reales* c_1 y c_2 tenemos

$$\iint_Q [c_1 s(x, y) + c_2 t(x, y)]\, dx\, dy = c_1 \iint_Q s(x, y)\, dx\, dy + c_2 \iint_Q t(x, y)\, dx\, dy .$$

TEOREMA 11.2. ADITIVIDAD. *Si Q está subdividido en dos rectángulos* Q_1 y Q_2,

$$\iint_Q s(x, y)\, dx\, dy = \iint_{Q_1} s(x, y)\, dx\, dy + \iint_{Q_2} s(x, y)\, dx\, dy .$$

TEOREMA 11.3. TEOREMA DE COMPARACIÓN. *Si $s(x, y) \leqslant t(x, y)$ para todo (x, y) de Q, entonces*

$$\iint_Q s(x, y)\, dx\, dy \leq \iint_Q t(x, y)\, dx\, dy .$$

En particular, si $t(x, y) \geqslant 0$ para todo (x, y) de Q, entonces

$$\iint_Q t(x, y)\, dx\, dy \geq 0 .$$

Dejamos como ejercicios las demostraciones de esos teoremas.

11.4 Definición de integral doble de una función definida y acotada en un rectángulo

Sea f una función definida y acotada en un rectángulo Q; en particular, supongamos que

$$|f(x, y)| \leq M \qquad \text{si} \quad (x, y) \in Q .$$

Entonces la función f puede quedar limitada por encima y por debajo por dos funciones escalonadas constantes s y t, siendo $s(x, y) = -M$ y $t(x, y) = M$ para todo (x, y) en Q. Consideremos ahora dos funciones escalonadas *cualesquiera* s y t, definidas en Q, tales que

(11.4) $$s(x, y) \leq f(x, y) \leq t(x, y) \quad \text{para todo punto } (x, y) \text{ de } Q.$$

DEFINICIÓN DE LA INTEGRAL DE UNA FUNCIÓN ACOTADA EN UN RECTÁNGULO. *Si existe un número I y solamente uno tal que*

(11.5) $$\iint_Q s \leq I \leq \iint_Q t$$

para todo par de funciones escalonadas que satisfagan las desigualdades (11.4), *dicho número I se llama integral doble de f extendida a Q y se representa por el símbolo*

$$\iint_Q f \qquad \text{o} \qquad \iint_Q f(x, y)\, dx\, dy .$$

Cuando existe tal número I se dice que la función f es integrable en Q.

11.5 Integrales dobles superior e inferior

La definición de integral doble es completamente análoga al caso uni-dimensional.

Supongamos que f es acotada en un rectángulo Q y sean s y t dos funciones escalonadas que satisfacen (11.4). Decimos que s está *por debajo de f*, y t está *por encima de f*, y escribimos $s \leqslant f \leqslant t$. Designemos con S el conjunto de todos los números $\iint_Q s$ obtenidos al tomar como s todas y cada una de las funciones

escalonadas por debajo de f, y sea T el conjunto de todos los números $\iint_Q t$ obtenidos al tomar como s todas y cada una de las funciones escalonadas por encima de f. Ambos conjuntos S y T son no vacíos puesto que f es acotada. Asimismo, $\iint_Q s \leq \iint_Q t$ si $s \leqslant f \leqslant t$, así que todo número de S es menor que todo número de T. Por consiguiente S tiene un extremo superior, y T tiene un extremo inferior, que satisfacen las desigualdades

$$\iint_Q s \leq \sup S \leq \inf T \leq \iint_Q t$$

para toda s y toda t que satisfagan $s \leqslant f \leqslant t$. Esto prueba que los dos números $\sup S$ e $\inf T$ satisfacen (11.5). Por lo tanto, f es integrable en Q si y sólo si $\sup S = \inf T$, en cuyo caso tenemos

$$\iint_Q f = \sup S = \inf T.$$

El número $\sup S$ se llama *integral inferior* de f y se representa por $\underline{I}(f)$. El número $\inf T$ es la *integral superior* de f y se representa por $\bar{I}(f)$. Así pues, tenemos

$$\underline{I}(f) = \sup \left\{ \iint_Q s \mid s \leq f \right\}, \qquad \bar{I}(f) = \inf \left\{ \iint_Q t \mid f \leq t \right\}.$$

El razonamiento anterior demuestra el siguiente teorema.

TEOREMA 11.4. *Si una función f acotada en un rectángulo Q tiene una integral inferior $\underline{I}(f)$ y una integral superior $\bar{I}(f)$ que satisfacen las desigualdades*

$$\iint_Q s \leq \underline{I}(f) \leq \bar{I}(f) \leq \iint_Q t$$

para todas las funciones escalonadas s y t que cumplen $s \leqslant f \leqslant t$, la función f es integrable en Q si y sólo si sus integrales superior e inferior son iguales, en cuyo caso tenemos

$$\iint_Q f = \underline{I}(f) = \bar{I}(f).$$

Puesto que las definiciones anteriores son del todo análogas a las del caso uni-dimensional, no debe sorprender que las propiedades de linealidad y aditividad y el teorema de comparación establecidos para funciones escalonadas en la sección 11.3, sean válidas también para las integrales dobles en general. Las demostraciones de esas propiedades son análogas a las del caso uni-dimensional y las omitiremos.

11.6 Cálculo de una integral doble por integración uni-dimensional reiterada

En la teoría de la integración unidimensional, el segundo teorema fundamental del cálculo proporciona un método para calcular integrales sin exigir la definición de integral en cada caso. El teorema siguiente logra el mismo resultado en dos dimensiones, y nos permite calcular ciertas integrales dobles mediante dos integraciones unidimensionales sucesivas. El resultado es una extensión de la fórmula (11.3), que ya hemos demostrado para funciones escalonadas.

TEOREMA 11.5. *Sea f una función definida y acotada en un rectángulo $Q = [a, b] \times [c, d]$, y supongamos que f es integrable en Q. Supongamos que para cada y fija en $[c, d]$ la integral uni-dimensional $\int_a^b f(x, y)\, dx$ exista, y designemos su valor por $A(y)$. Si existe la integral $\int_c^d A(y)$, es igual a la integral doble $\iint_Q f$. Es decir, tenemos la fórmula*

$$\iint_Q f(x, y)\, dx\, dy = \int_c^d \left[\int_a^b f(x, y)\, dx\right] dy. \tag{11.6}$$

Demostración. Elijamos dos funciones escalonadas cualesquiera s y t que satisfagan $s \leqslant f \leqslant t$ en Q. Integrando con respecto a x en el intervalo $[a, b]$ tenemos

$$\int_a^b s(x, y)\, dx \le A(y) \le \int_a^b t(x, y)\, dx .$$

Puesto que la integral $\int_c^d A(y)\, dy$ existe, podemos integrar esas dos desigualdades con respecto a y en $[c, d]$ y aplicar la ecuación (11.3) obteniendo

$$\iint_Q s \le \int_c^d A(y)\, dy \le \iint_Q t .$$

Por consiguiente $\int_c^d A(y)\, dy$ es un número comprendido entre $\iint_Q s$ e $\iint_Q t$ para todas las funciones s y t que aproximan f por debajo y por encima, respectivamente. Puesto que f es integrable en Q, el único número con esa propiedad es la

integral doble de f en Q. Por tanto $\int_c^d A(y)\,dy = \iint_Q f$, lo que demuestra la ecuación (11.6).

Se dice que la fórmula (11.6) proporciona una evaluación de la integral doble mediante integración reiterada o repetida. El proceso consiste pues en integrar f respecto a x entre a y b (dejando fija y), y luego se integra el resultado respecto a y entre c y d. Si invertimos el orden de integración, obtenemos una fórmula parecida, a saber,

(11.7)
$$\iint_Q f(x, y)\,dx\,dy = \int_a^b \left[\int_c^d f(x, y)\,dy\right] dx,$$

que es válida si suponemos que $\int_c^d f(x, y)\,dy$ existe para cada x fija en $[a, b]$ y que es integrable en $[a, b]$.

11.7 Interpretación geométrica de la integral doble como un volumen

El teorema 11.5 tiene una sencilla interpretación geométrica, representada en la figura 11.4. Si f es no negativa, el conjunto S de los puntos (x, y, z) en el espacio de tres dimensiones tales que (x, y) está en Q y $0 \leq z \leq f(x, y)$ se llama

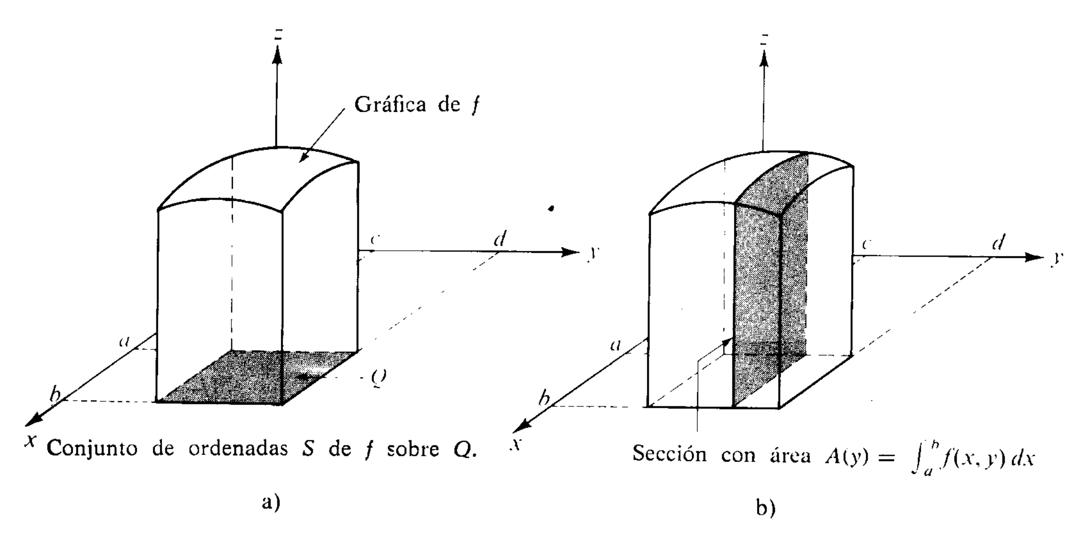


FIGURA 11.4 *El volumen de S es la integral del área de la sección:*
$$v(S) = \int_c^d A(y)\,dy.$$

conjunto de ordenadas de f sobre Q. Consta de los puntos comprendidos entre el rectángulo Q y la superficie $z = f(x, y)$. (Véase figura 11.4 a).) Para cada y en el intervalo $[c, d]$, la integral

$$A(y) = \int_a^b f(x, y)\, dx$$

es el área de la sección producida en S por un plano paralelo al plano xz (la región sombreada en la figura 11.4 b)). Puesto que el área de la sección $A(y)$ es integrable en $[c, d]$, el teorema 2.7 del volumen I nos dice que la integral $\int_c^d A(y)\, dy$ es igual a $v(S)$, el volumen de S. Así pues, para integrandos no negativos, el teorema 11.5 muestra que el volumen del conjunto de ordenadas de f sobre Q es igual a la integral doble $\iint\limits_Q f$.

La fórmula (11.7) nos da otro camino para calcular el volumen del conjunto de ordenadas. En ella integramos el área de las secciones producidas por planos paralelos al plano yz.

11.8 Ejemplos resueltos

En esta sección ilustramos el teorema 11.5 con dos ejemplos numéricos.

EJEMPLO 1. Si $Q = [-1, 1] \times [0, \pi/2]$, calcular $\iint\limits_Q (x \operatorname{sen} y - ye^x)\, dx\, dy$, en el supuesto de que la integral exista. La región de integración está representada en la figura 11.5.

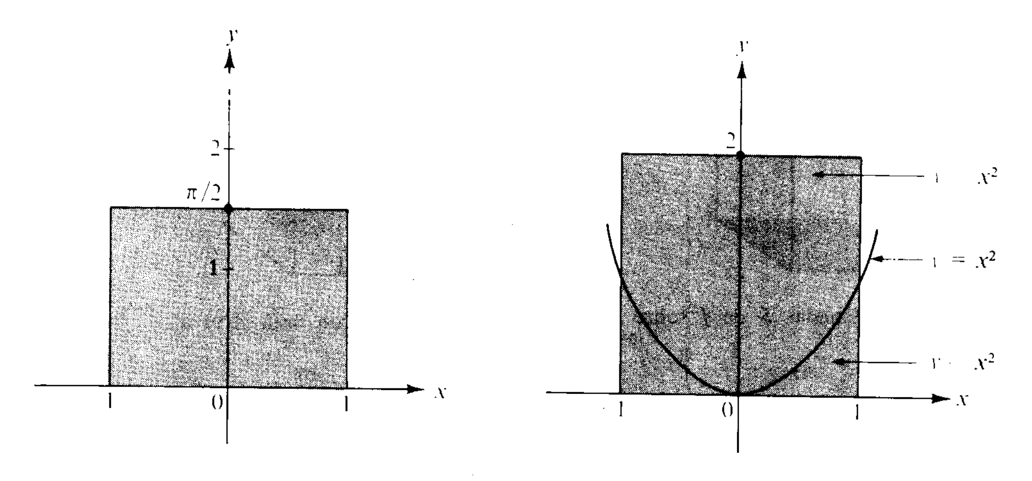


FIGURA 11.5 *Región de integración para el ejemplo 1.*

FIGURA 11.6 *Región de integración para el ejemplo 2.*

Solución. Integrando primero respecto a x y llamando al resultado $A(y)$, tenemos

$$A(y) = \int_{-1}^{1} (x \operatorname{sen} y - ye^x)\, dx = \left(\frac{x^2}{2} \operatorname{sen} y - ye^x\right)\Bigg|_{x=-1}^{x=1} = -ey + y/e\,.$$

Aplicando el teorema 11.5 encontramos

$$\iint_Q (x \operatorname{sen} y - ye^x)\, dx\, dy = \int_0^{\pi/2} A(y)\, dy = \int_0^{\pi/2} (-ey + y/e)\, dy$$

$$= (1/e - e) \int_0^{\pi/2} y\, dy = (1/e - e)\pi^2/8\,.$$

Como comprobación del cálculo podemos integrar primero respecto a y:

$$\iint_Q (x \operatorname{sen} y - ye^x)\, dx\, dy = \int_{-1}^{1} \left[\int_0^{\pi/2} (x \operatorname{sen} y - ye^x)\, dy\right] dx$$

$$= \int_{-1}^{1} (-x \cos y - \tfrac{1}{2} y^2 e^x)\Big|_{y=0}^{y=\pi/2} dx$$

$$= \int_{-1}^{1} (-\pi^2 e^x/8 + x)\, dx = (1/e - e)\pi^2/8\,.$$

EJEMPLO 2. Si $Q = [-1, 1] \times [0, 2]$, calcular la integral doble $\iint_Q \sqrt{|y - x^2|}\, dx\, dy$, supuesta su existencia.

Solución. Si integramos primero respecto a y y el resultado es $H(x)$, tenemos $H(x) = \int_0^2 \sqrt{|y - x^2|}\, dy$. La región de integración es el rectángulo de la figura 11.6. La parábola $y = x^2$ está allí dibujada debido a la presencia de $|y - x^2|$en el integrando. Por encima de esa parábola se tiene $y > x^2$ y por debajo $y < x^2$. Esto nos sugiere que escindamos la integral de $H(x)$ del modo siguiente:

$$H(x) = \int_0^2 \sqrt{|y - x^2|}\, dy = \int_0^{x^2} \sqrt{x^2 - y}\, dy + \int_{x^2}^{2} \sqrt{y - x^2}\, dy\,.$$

Recordemos que x se considera constante en cada una de esas integrales. En la primera efectuamos el cambio de variable $t = x^2 - y$ y en la segunda $t = y - x^2$. Esto nos da

$$H(x) = \int_0^2 \sqrt{|y - x^2|}\, dy = -\int_{x^2}^{0} \sqrt{t}\, dt + \int_0^{2-x^2} \sqrt{t}\, dt = \tfrac{2}{3}x^3 + \tfrac{2}{3}(2 - x^2)^{3/2}.$$

Aplicando el teorema 11.5 encontramos

$$\iint_Q \sqrt{|y - x^2|}\, dx\, dy = \int_{-1}^{1} [\tfrac{2}{3} x^3 + \tfrac{2}{3}(2 - x^2)^{3/2}]\, dx = \tfrac{4}{3}\int_0^1 (2 - x^2)^{3/2}\, dx$$

$$= \frac{1}{3}\left[x(2 - x^2)^{3/2} + 3x\sqrt{2 - x^2} + 6 \operatorname{arc\,sen}\left(\frac{x}{\sqrt{2}}\right)\right]\Bigg|_0^1 = \frac{4}{3} + \frac{\pi}{2}.$$

Análogo resultado se consigue invirtiendo el orden de integración, pero los cálculos son más complicados.

11.9 Ejercicios

Calcular las integrales dobles por integración sucesiva en los ejercicios del 1 al 8, supuesta la existencia de cada integral.

1. $\iint_Q xy(x + y)\, dx\, dy$, donde $Q = [0, 1] \times [0, 1]$.

2. $\iint_Q (x^3 + 3x^2y + y^3)\, dx\, dy$, donde $Q = [0, 1] \times [0, 1]$.

3. $\iint_Q (\sqrt{y} + x - 3xy^2)\, dx\, dy$, donde $Q = [0, 1] \times [1, 3]$.

4. $\iint_Q \operatorname{sen}^2 x \operatorname{sen}^2 y\, dx\, dy$, donde $Q = [0, \pi] \times [0, \pi]$.

5. $\iint_Q \operatorname{sen}(x + y)\, dx\, dy$, donde $Q = [0, \pi/2] \times [0, \pi/2]$.

6. $\iint_Q |\cos(x + y)|\, dx\, dy$, donde $Q = [0, \pi] \times [0, \pi]$.

7. $\iint_Q f(x + y)\, dx\, dy$, donde $Q = [0, 2] \times [0, 2]$, y $f(t)$ representa el mayor entero $\leq t$.

8. $\iint_Q y^{-3} e^{tx/y}\, dx\, dy$, donde $Q = [0, t] \times [1, t]$, $t > 0$.

9. Si Q es un rectángulo, demostrar que una integral doble de la forma $\iint_Q f(x)g(y)\, dx\, dy$ es igual al producto de dos integrales unidimensionales. Establecer las hipótesis relativas a la existencia.

10. Sea f una función definida en el rectángulo $Q = [0, 1] \times [0, 1]$ del siguiente modo:

$$f(x, y) = \begin{cases} 1 - x - y & \text{si } x + y \leq 1, \\ 0 & \text{en los demás puntos de } Q. \end{cases}$$

Representar el conjunto de ordenadas de f sobre Q y calcular su volumen por doble integración (Supóngase la existencia de la integral.)

11. Resolver el ejercicio 10 cuando

$$f(x, y) = \begin{cases} x + y & \text{si } x^2 \leq y \leq 2x^2, \\ 0 & \text{en los demás puntos de } Q. \end{cases}$$

12. Resolver el ejercicio 10 cuando $Q = [-1, 1] \times [-1, 1]$ y

$$f(x, y) = \begin{cases} x^2 + y^2 & \text{si } x^2 + y^2 \leq 1, \\ 0 & \text{en los demás puntos de } Q. \end{cases}$$

13. Sea f una función definida en el rectángulo $Q = [1, 2] \times [1, 4]$ del siguiente modo:

$$f(x, y) = \begin{cases} (x + y)^{-2} & \text{si } x \leq y \leq 2x, \\ 0 & \text{en los demás puntos de } Q. \end{cases}$$

Indicar, mediante un dibujo, la porción de Q en la que f es no nula y calcular el valor de la integral doble $\iint_Q f$, supuesta su existencia.

14. Sea f una función definida en el rectángulo $Q = [0, 1] \times [0, 1]$ del siguiente modo:

$$f(x, y) = \begin{cases} 1 & \text{si } x = y, \\ 0 & \text{si } x \neq y. \end{cases}$$

Demostrar la existencia de la integral doble $\iint_Q f$ y que es igual a 0.

11.10 Integrabilidad de funciones continuas

El teorema de la continuidad uniforme (teorema 9.10) puede aplicarse para demostrar la integrabilidad de una función continua en un rectángulo.

TEOREMA 11.6. INTEGRABILIDAD DE FUNCIONES CONTINUAS. *Si una fun-*

ción f es continua en un rectángulo $Q = [a, b] \times [c, d]$, entonces f es integrable en Q. Además, el valor de la integral puede obtenerse por integración sucesiva,

$$\iint_Q f = \int_c^d \left[\int_a^b f(x, y)\, dx\right] dy = \int_a^b \left[\int_c^d f(x, y)\, dy\right] dx. \tag{11.8}$$

Demostración. El teorema 9.8 demuestra que f es acotada en Q, por lo que f tiene una integral superior y una integral inferior. Demostraremos que $\underline{I}(f) = \bar{I}(f)$. Elijamos $\epsilon > 0$. Según el teorema de la continuidad uniforme, para ese ϵ existe una partición P de Q en un número finito (por ejemplo n) de subrectángulos $Q_1, \ldots, Q_n$ tal que la oscilación de f en todo subrectángulo es menor que ϵ. Designemos con $M_k(f)$ y $m_k(f)$, respectivamente, los valores máximo y mínimo absolutos de f en Q_k. Tenemos entonces

$$M_k(f) - m_k(f) < \epsilon$$

para cada $k = 1, 2, \ldots, n$. Sean ahora s y t dos funciones escalonadas definidas en el interior de cada Q_k así:

$$s(x) = m_k(f), \qquad t(x) = M_k(f) \qquad \text{si} \quad x \in \operatorname{int} Q_k.$$

En los puntos frontera definamos $s(x) = m$ y $t(x) = M$, siendo m y M, respectivamente, los valores mínimo y máximo absolutos de f en Q. Tenemos entonces $s \leqslant f \leqslant t$ para todo x de Q. Asimismo, tenemos

$$\iint_Q s = \sum_{k=1}^n m_k(f) a(Q_k) \qquad \text{y} \qquad \iint_Q t = \sum_{k=1}^n M_k(f) a(Q_k),$$

en donde $a(Q_k)$ es el área del rectángulo Q_k. La diferencia de esas dos integrales es

$$\iint_Q t - \iint_Q s = \sum_{k=1}^n \{M_k(f) - m_k(f)\} a(Q_k) < \epsilon \sum_{k=1}^n a(Q_k) = \epsilon a(Q),$$

en donde $a(Q)$ es el área de Q. Puesto que $\iint_Q s \leq \underline{I}(f) \leq \bar{I}(f) \leq \iint_Q t$ obtenemos la desigualdad

$$0 \leq \bar{I}(f) - \underline{I}(f) \leq \epsilon a(Q).$$

Haciendo que $\epsilon \to 0$ vemos que $\underline{I}(f) = \bar{I}(f)$, por lo que f es integrable en Q.

Demostramos seguidamente que la integral doble es igual a la integral reiterada de (11.8). Para cada y fija de $[c, d]$ la integral unidimensional $\int_a^b f(x, y)\, dx$ existe ya que el integrando es continuo en Q. Sea $A(y) = \int_a^b f(x, y) dx$. Demostraremos que A es continua en $[c, d]$. Si y e y_1 son dos puntos cualesquiera de $[c, d]$ tenemos

$$A(y) - A(y_1) = \int_a^b \{f(x, y) - f(x, y_1)\}\, dx$$

de la que obtenemos

$$|A(y) - A(y_1)| \leq (b - a) \max_{a \leq x \leq b} |f(x, y) - f(x, y_1)| = (b - a)\, |f(x_1, y) - f(x_1, y_1)|$$

en donde x_1 es un punto de $[a, b]$ en el que $|f(x, y) - f(x, y_1)|$ alcanza su máximo. Esta desigualdad demuestra que $A(y) \to A(y_1)$ cuando $y \to y_1$, así que A es continua en y_1. Por consiguiente existe la integral $\int_c^d A(y)\, dy$ y, según el teorema 11.5, es igual a $\iint_Q f$. Se sigue un razonamiento parecido cuando el orden de integración se invierte.

11.11 Integrabilidad de funciones acotadas con discontinuidades

Sea f una función definida y acotada en un rectángulo Q. En el teorema 11.6 se demostró que la integral doble de f sobre Q existe si f es continua en todos los puntos de Q. En esta sección demostramos que también existe la integral si f tiene discontinuidades en Q, con tal que el conjunto de las discontinuidades no sea demasiado grande. Para medir el tamaño del conjunto de discontinuidades introducimos el concepto siguiente.

DEFINICIÓN DE CONJUNTO ACOTADO DE CONTENIDO NULO. *Sea A un subconjunto acotado del plano. Se dice que el conjunto A tiene contenido nulo si para todo $\epsilon > 0$ existe un conjunto finito de rectángulos cuya reunión contiene A y cuya suma de áreas no supera ϵ.*

Es decir, un conjunto plano acotado de contenido nulo puede cubrirse con una reunión de rectángulos cuya área total sea tan pequeña como se quiera.

Las siguientes proposiciones relativas a los conjuntos acotados de contenido nulo son sencillas consecuencias de esa definición. Dejamos al lector las demostraciones.

a) Cualquier conjunto finito de puntos del plano tiene contenido nulo.

b) La reunión de un número finito de conjuntos acotados de contenido nulo también es de contenido nulo.

c) Todo subconjunto de un conjunto de contenido nulo tiene contenido nulo.

d) Todo segmento de recta tiene contenido nulo.

TEOREMA 11.7. *Sea f una función definida y acotada en un rectángulo $Q = [a, b] \times [c, d]$. Si el conjunto de discontinuidades de f en Q es un conjunto de contenido nulo existe la integral doble $\iint_Q f$.*

Demostración. Sea $M > 0$ tal que $|f| \leqslant M$ en Q. Llamemos D al conjunto de discontinuidades de f en Q. Dado $\delta > 0$, sea P una partición de Q tal que la suma de las áreas de todos los subrectángulos de P que contienen puntos de D sea menor que δ. (Estò es posible puesto que D tiene contenido nulo.) Definamos en esos subrectángulos las funciones escalonadas s y t como sigue:

$$s(x) = -M, \qquad t(x) = M.$$

En los subrectángulos de P restantes definamos s y t como se hizo en la demostración del teorema 11.6. Tenemos entonces $s \leqslant f \leqslant t$ en todo Q. Razonando como en la demostración del teorema 11.6 obtenemos la desigualdad

$$\iint_Q t - \iint_Q s \leq \epsilon a(Q) + 2M\delta. \tag{11.9}$$

El primer término, $\epsilon a(Q)$, procede del cálculo de la integral de $t - s$ en los subrectángulos que sólo contienen puntos de continuidad de f; el segundo término, $2M\delta$, procede de la integral de $t - s$ en los subrectángulos que contienen puntos de D. De (11.9) obtenemos la desigualdad

$$0 \leq \bar{I}(f) - \underline{I}(f) \leq \epsilon a(Q) + 2M\delta.$$

Haciendo que $\epsilon \to 0$ encontramos $0 \leqslant \bar{I}(f) - \underline{I}(f) \leqslant 2M\delta$. Puesto que δ es arbitrario esto implica que $\bar{I}(f) = \underline{I}(f)$, así que f es integrable en Q.

11.12 Integrales dobles extendidas a regiones más generales

Hasta aquí la integral doble sólo se ha definido para regiones de integración rectangulares. No obstante, no hay dificultad para extender el concepto a regiones más generales. Sea S una región acotada, e incluyamos S en un rectángulo Q. Sea f una función definida y acotada en S. Definamos una nueva función $\tilde{f}$ en Q del siguiente modo:

$$\tilde{f}(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in S, \\ 0 & \text{si } (x, y) \in Q - S. \end{cases} \tag{11.10}$$

Es decir, extendemos la definición de f a todo el rectángulo Q haciendo que la función valga 0 fuera de S. Nos preguntamos ahora si la función $\tilde{f}$ es o no integrable en Q. Si lo es, decimos que f es integrable en S y que, *por definición,*

$$\iint_S f = \iint_Q \tilde{f}.$$

Consideremos primero conjuntos S de puntos del plano xy definidos así:

$$S = \{(x, y) \mid a \leq x \leq b \quad \text{y} \quad \varphi_1(x) \leq y \leq \varphi_2(x)\},$$

en donde φ_1 y φ_2 son funciones continuas en un intervalo cerrado $[a, b]$ que satisfacen $\varphi_1 \leqslant \varphi_2$. Ejemplo de una tal región, que llamamos región del tipo *I*, es el de la figura 11.7. En una región del tipo I, para cada punto t de $[a, b]$ la recta vertical $x = t$ corta a S en un segmento que une la curva $y = \varphi_1(x)$ a $y = \varphi_2(x)$. Una tal región es acotada porque φ_1 y φ_2 son continuas y por tanto acotadas en $[a, b]$.

Otro tipo de región T (tipo *II*) puede definirse así:

$$T = \{(x, y) \mid c \leq y \leq d \quad \text{y} \quad \psi_1(y) \leq x \leq \psi_2(y)\},$$

en donde ψ_1 y ψ_2 son continuas en un intervalo $[c, d]$ siendo $\psi_1 \leqslant \psi_2$. En la figura 11.8 se muestra un ejemplo. En este caso las rectas horizontales cortan T en

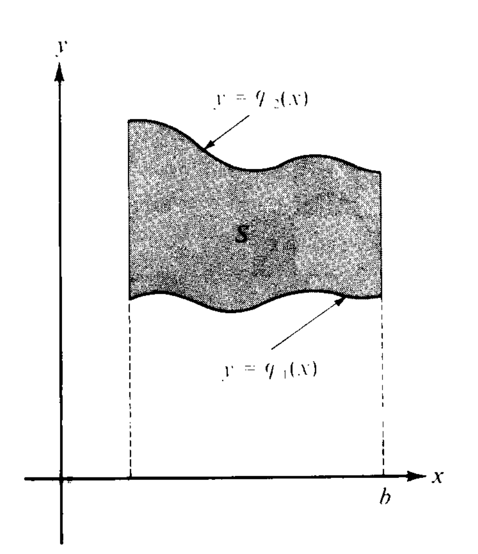


Figura 11.7 *Región S de tipo I.*

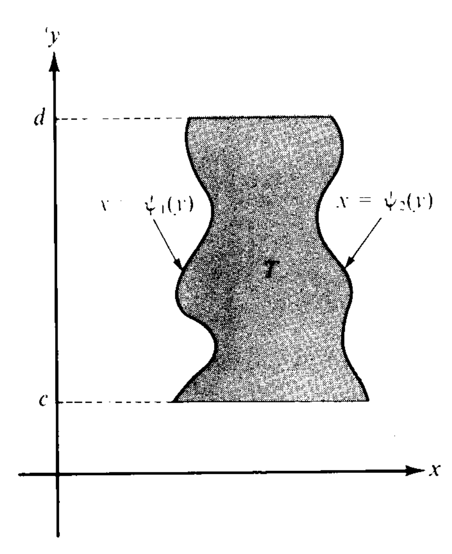


Figura 11.8 *Región T de tipo II.*

segmentos que unen las curvas $x_1 = \psi_1(y)$ y $x_2 = \psi_2(y)$. Las regiones del tipo II también son acotadas. Todas las regiones que consideraremos serán de uno u otro de esos dos tipos o podrán descomponerse en un número finito de fragmentos, cada uno de los cuales será de uno de esos dos tipos.

Sea f una función definida y acotada en una región S de tipo I. Incluyamos S en un rectángulo Q y definamos $\tilde{f}$ en Q como está indicado en la ecuación (11.10). Las discontinuidades de $\tilde{f}$ en Q serán las de f en S y además aquellos puntos de la frontera de S en los que f no es cero. La frontera de S está constituida por las gráficas de φ_1, de φ_2 y por los dos segmentos de recta verticales que unen los extremos de las gráficas. Cada uno de esos segmentos de recta tiene contenido nulo. El teorema que sigue demuestra que cada una de las gráficas también tiene contenido nulo.

TEOREMA 11.8. *Si φ es una función real continua en un intervalo $[a, b]$, entonces la gráfica de φ tiene contenido nulo.*

Demostración. Designemos con A la gráfica de φ, esto es, pongamos

$$A = \{(x, y) \mid y = \varphi(x) \quad \text{y} \quad a \le x \le b\}.$$

Elijamos cualquier $\epsilon > 0$. Apliquemos el teorema de la continuidad uniforme (teorema 3.13 del volumen I) para obtener una partición P de $[a, b]$ en un número finito de subintervalos tales que la oscilación de φ en todo subintervalo de P sea menor que $\epsilon/(b - a)$ Por consiguiente, en cada subintervalo de P la gráfica de φ está dentro de un rectángulo de altura $\epsilon/(b - a)$. Luego toda la grá-

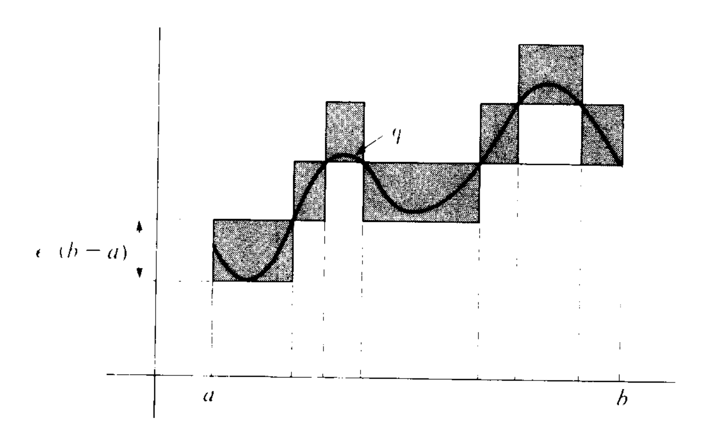


FIGURA 11.9 *Demostración de que la gráfica de una función continua φ tiene contenido nulo.*

fica de φ queda incluida en una reunión finita de rectángulos, cuya suma de áreas es ϵ. (Véase la figura 11.9.) Esto demuestra que la gráfica de φ tiene contenido nulo.

El teorema siguiente demuestra que la integral doble $\iint_S f$ existe si f es continua en int S, (el interior de S). Éste es el conjunto

$$\text{int } S = \{(x, y) \mid a < x < b \quad \text{y} \quad \varphi_1(x) < y < \varphi_2(x)\}.$$

TEOREMA 11.9. *Sea S una región del tipo* I, *comprendida entre las gráficas de* φ_1 *y* φ_2. *Supongamos que f está definida y es acotada en S y que es continua en el interior de S. Existe entonces la integral doble* $\iint_S f$ *y puede calcularse mediante integración uni-dimensional reiterada.*

(11.11) $$\iint_S f(x, y)\, dx\, dy = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\, dy\right] dx.$$

Demostración. Sea $Q = [a, b] \times [c, d]$ un rectángulo que contiene S, y sea $\tilde{f}$ la función definida en (11.10). Los únicos puntos que pueden ser de discontinuidad para $\tilde{f}$ son los de la frontera de S. Puesto que la frontera de S tiene contenido nulo, $\tilde{f}$ es integrable en Q. Para cada x de (a, b) fijo existe la integral unidimensional $\int_c^d \tilde{f}(x, y)\, dy$ puesto que el integrando tiene a lo más dos discontinuidades en $[c, d]$. Aplicando la versión del teorema 11.5 dada en la ecuación (11.7) tenemos

(11.12) $$\iint_Q \tilde{f} = \int_a^b \left[\int_c^d \tilde{f}(x, y)\, dy\right] dx.$$

Si $\varphi_1(x) \leqslant y \leqslant \varphi_2(x)$, $\tilde{f}(x,y) = f(x, y)$, y para los demás valores de y en $[c, d]$, $\tilde{f}(x, y) = 0$. Por consiguiente

$$\int_c^d \tilde{f}(x, y)\, dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\, dy$$

por lo cual la ecuación (11.12) implica (11.11).

Existe, naturalmente, un teorema análogo para una región T del tipo II. Si f está definida y es acotada en T y es continua en el interior de T, f es integrable en T y la fórmula para la integración reiterada es ahora

(11.13) $$\iint_T f(x, y)\, dx\, dy = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\, dx\right] dy.$$

Ciertas regiones son del tipo I y del tipo II a la vez. (Por ejemplo regiones

limitadas por circunferencias y elipses.) En estos casos el orden de integración es indiferente y podemos escribir

$$\int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\, dy\right] dx = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\, dx\right] dy.$$

En algunos casos una de esas integrales puede ser de cálculo más sencillo que la otra; es conveniente examinarlas antes de intentar el cálculo.

11.13 Aplicaciones a áreas y volúmenes

Sea S una región del tipo I dada por

$$S = \{(x, y) \mid a \leq x \leq b \quad \text{y} \quad \varphi_1(x) \leq y \leq \varphi_2(x)\}.$$

Aplicando el teorema 11.9 poniendo $f(x, y) = 1$ para todo (x, y) de S obtenemos

$$\iint_S dx\, dy = \int_a^b [\varphi_2(x) - \varphi_1(x)]\, dx.$$

Según el teorema 2.1 del volumen I, la integral del segundo miembro es igual al área de la región S. Así pues, las integrales dobles pueden emplearse para calcular áreas.

Si f es no negativa, el conjunto de puntos (x, y, z) del espacio de tres dimensiones tales que $(x, y) \in S$ y $0 \leqslant z \leqslant f(x, y)$ se denomina *conjunto de ordenadas de f sobre S*. En la figura 11.10 se ve un ejemplo. Si f es no negativa y continua en S, la integral

$$\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\, dy$$

representa el área de una sección del conjunto de ordenadas producida al cortarlo por un plano paralelo al plano yz (región sombreada de la figura 11.10). La fórmula (11.11) del teorema 11.9 demuestra que la integral doble de f sobre S es igual a la integral del área de las secciones. Luego la integral doble $\iint_S f$ es igual al volumen del conjunto de ordenadas de f sobre S. (Véase el teorema 2.7 del volumen I, pág. 139.)

En general, si f y g son ambas continuas en S siendo $f \leqslant g$, la integral doble $\iint_S (g - f)$ es igual al volumen del sólido comprendido entre las gráficas de las funciones f y g. A las regiones del tipo II se aplican observaciones análogas.

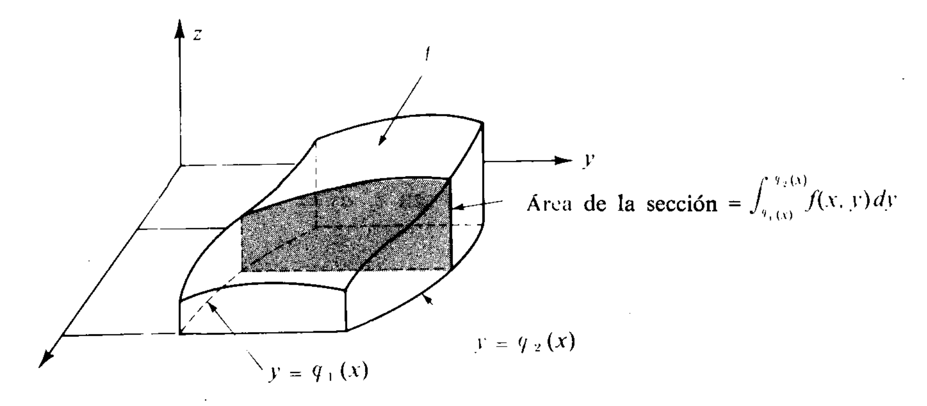


FIGURA 11.10 *La integral* $\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\,dy$ *es el área de la sección del conjunto de ordenadas. La integral reiterada* $\int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\,dy\right] dx$ *es el volumen del conjunto de ordenadas.*

11.14 Ejemplos resueltos

EJEMPLO 1. Calcular el volumen del sólido limitado por el elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 .$$

Solución. El sólido en cuestión está comprendido entre las gráficas de las dos funciones f y g, siendo

$$g(x, y) = c\sqrt{1 - x^2/a^2 - y^2/b^2} \qquad f(x, y) = -g(x, y).$$

Aquí $(x, y) \in S$, siendo S la región elíptica dada por

$$S = \left\{(x, y) \,\middle|\, \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\}.$$

Aplicando el teorema 11.9 y teniendo en cuenta la simetría encontramos que el volumen V del sólido elipsoidal viene dado por

$$V = \iint_S (g - f) = 2\iint_S g = 8c \int_0^a \left[\int_0^{b\sqrt{1-x^2/a^2}} \sqrt{1 - x^2/a^2 - y^2/b^2}\,dy\right] dx .$$

Pongamos $A = \sqrt{1 - x^2/a^2}$. La integral interior es

$$\int_0^{bA} \sqrt{A^2 - y^2/b^2}\, dy = A \int_0^{bA} \sqrt{1 - y^2/(Ab)^2}\, dy .$$

Mediante el cambio de variable $y = Ab$ sen t, $dy = Ab \cos t\, dt$, encontramos que la última integral es igual a

$$A^2b \int_0^{\pi/2} \cos^2 t\, dt = \frac{\pi}{4} A^2 b = \frac{\pi b}{4}\left(1 - \frac{x^2}{a^2}\right).$$

Por consiguiente

$$V = 8c \int_0^a \frac{\pi b}{4}\left(1 - \frac{x^2}{a^2}\right) dx = \frac{4}{3} \pi abc .$$

En el caso particular $a = b = c$ el sólido es una esfera de radio a y el volumen es $\frac{4}{3}\pi a^3$.

EJEMPLO 2. La integral doble de una función positiva f, $\iint_S f(x, y)\, dx\, dy$, se reduce a la integral reiterada

$$\int_0^1 \left[\int_{x^2}^{x} f(x, y)\, dy\right] dx .$$

Determinar la región S e intercambiar el orden de integración.

Solución. Para cada x fijada entre 0 y 1, la integración con respecto a y se efectúa en el intervalo entre x^2 y x. Esto significa que la región es del tipo I y está limitada entre los dos curvas $y = x^2$ e $y = x$. La región S es el conjunto de puntos situados entre las dos curvas y sobre el intervalo [0, 1]. (Véase figura 11.11) Puesto que S es también del tipo II podemos invertir el orden de integración obteniendo

$$\int_0^1 \left[\int_y^{\sqrt{y}} f(x, y)\, dx\right] dy .$$

EJEMPLO 3. Una integral doble de una función f, $\iint_S f(x, y)\, dx\, dy$, se reduce a la integral reiterada:

$$\int_0^3 \left[\int_{4y/3}^{\sqrt{25-y^2}} f(x, y)\, dx\right] dy .$$

Determinar la región S e invertir el orden de integración.

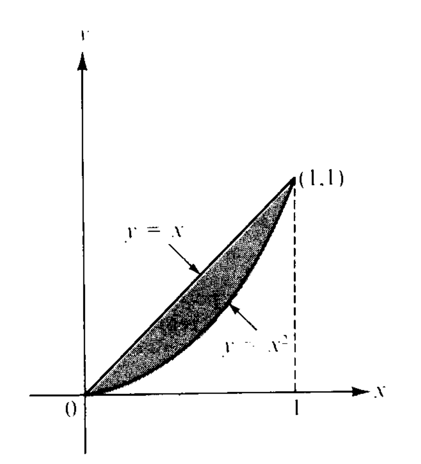


FIGURA 11.11 *Ejemplo 2.*

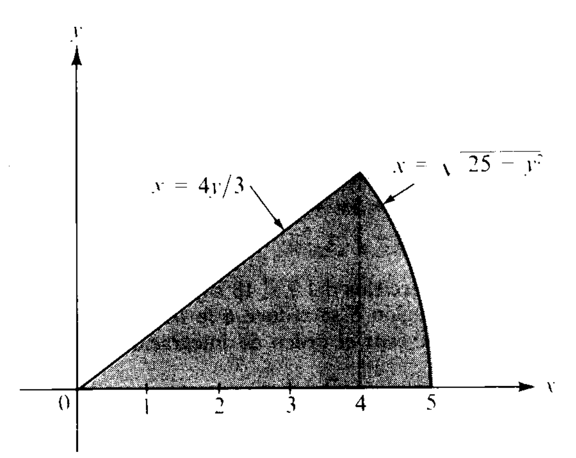


FIGURA 11.12 *Ejemplo 3.*

Solución. Para cada y fija entre 0 y 3, la integración respecto a x se efectúa en el intervalo entre $4y/3$ y $\sqrt{25-y^2}$. Por lo tanto la región es del tipo II y está limitada por las dos curvas $x = 4y/3$ y $x = \sqrt{25-y^2}$. Esta región, representada en la figura 11.12, es un sector circular. Cuando el orden de integración se invierte, la región se escinde en dos regiones del tipo I; el resultado es la suma de dos integrales:

$$\int_0^4 \left[\int_0^{3x/4} f(x, y)\, dy\right] dx + \int_4^5 \left[\int_0^{\sqrt{25-x^2}} f(x, y)\, dy\right] dx .$$

11.15 Ejercicios

En los ejercicios del 1 al 5, dibujar la región de integración y calcular la integral doble.

1. $\iint_S x \cos(x + y)\, dx\, dy$, siendo S el triángulo de vértices $(0, 0)$, $(\pi, 0)$, (π, π).

2. $\iint_S (1 + x) \operatorname{sen} y\, dx\, dy$, siendo S el trapezoide de vértices $(0, 0)$, $(1, 0)$, $(1, 2)$, $(0, 1)$.

3. $\iint_S e^{x+y}\, dx\, dy$, siendo $S = \{(x, y) \mid |x| + |y| \leq 1\}$.

4. $\iint_S x^2y^2\, dx\, dy$, siendo S la porción acotada del primer cuadrante situada entre las dos hipérbolas $xy = 1$ y $xy = 2$ y las líneas rectas $y = x$ e $y = 4x$.

5. $\iint_S (x^2 - y^2)\, dx\, dy$, siendo S la región limitada por $y = \operatorname{sen} x$ y el intervalo $[0, \pi]$.

6. Una pirámide está limitada por los tres planos coordenados y el plano $x + 2y + 3z = 6$. Representar el sólido y calcular su volumen por doble integración.
7. Un sólido está limitado por la superficie $z = x^2 - y^2$, el plano xy, y los planos $x = 1$ y $x = 3$. Representar el sólido y calcular su volumen por doble integración.
8. Calcular por doble integración, el volumen del conjunto de ordenadas de f sobre S si:

a) $f(x, y) = x^2 + y^2$ y $S = \{(x, y) \mid |x| \leq 1, |y| \leq 1\}$.
b) $f(x, y) = 3x + y$ y $S = \{(x, y) \mid 4x^2 + 9y^2 \leq 36, x > 0, y > 0\}$.
c) $f(x, y) = y + 2x + 20$ y $S = \{(x, y) \mid x^2 + y^2 \leq 16\}$.

En los ejercicios del 9 al 18 suponer que la integral doble de una función positiva f extendida a una región S se reduce a la integral reiterada que se da. En cada caso, representar la región S e invertir el orden de integración.

9. $\int_0^1 \left[\int_0^y f(x, y)\, dx\right] dy$.
10. $\int_0^2 \left[\int_{y^2}^{2y} f(x, y)\, dx\right] dy$.
11. $\int_1^4 \left[\int_{\sqrt{x}}^{2} f(x, y)\, dy\right] dx$.
12. $\int_1^2 \left[\int_{2-x}^{\sqrt{2x-x^2}} f(x, y)\, dy\right] dx$.
13. $\int_{-6}^2 \left[\int_{(x^2-4)/4}^{2-x} f(x, y)\, dy\right] dx$.
14. $\int_1^e \left[\int_0^{\log x} f(x, y)\, dy\right] dx$.
15. $\int_{-1}^1 \left[\int_{-\sqrt{1-x^2}}^{1-x^2} f(x, y)\, dy\right] dx$.
16. $\int_0^1 \left[\int_{x^3}^{x^2} f(x, y)\, dy\right] dx$.
17. $\int_0^\pi \left[\int_{-\operatorname{sen}(x/2)}^{\operatorname{sen} x} f(x, y)\, dy\right] dx$.
18. $\int_0^4 \left[\int_{-\sqrt{4-y}}^{(y-4)/2} f(x, y)\, dx\right] dy$.

19. Al calcular por doble integración el volumen V situado debajo del paraboloide $z = x^2 + y^2$ y limitado inferiormente por una cierta región S del plano xy, se ha llegado a la siguiente suma de integrales reiteradas:

$$V = \int_0^1 \left[\int_0^y (x^2 + y^2)\, dx\right] dy + \int_1^2 \left[\int_0^{2-y} (x^2 + y^2)\, dx\right] dy.$$

Dibujar la región S y expresar V mediante una integral reiterada en la que el orden de integración esté invertido. Efectuar, también, la integración y calcular V.

20. Al calcular por doble integración el volumen V limitado por encima por la superficie $z = f(x, y)$ y por la parte inferior por una cierta región S del plano xy, se ha llegado a la siguiente suma de integrales reiteradas:

$$V = \int_0^{a \operatorname{sen} c} \left[\int_{\sqrt{a^2-y^2}}^{\sqrt{b^2-y^2}} f(x, y)\, dx\right] dy + \int_{a \operatorname{sen} c}^{b \operatorname{sen} c} \left[\int_{y \cot c}^{\sqrt{b^2-y^2}} f(x, y)\, dx\right] dy.$$

Siendo $0 < a < b$ y $0 < c < \pi/2$, dibujar la región S dando las ecuaciones de todas las curvas que constituyen su frontera.

21. Al calcular por doble integración el volumen V limitado por encima por la superficie

$z = f(x, y)$ y por la parte inferior por una cierta región S del plano xy, se ha llegado a la siguiente suma de integrales reiteradas:

$$V = \int_1^2 \left[\int_x^{x^3} f(x, y)\, dy \right] dx + \int_2^8 \left[\int_x^8 f(x, y)\, dy \right] dx .$$

a) Dibujar la región S y expresar el volumen V mediante una integral reiterada en la que el orden de integración esté invertido.

b) Efectuar la integración y calcular V cuando $f(x, y) = e^x (x/y)^{1/2}$.

22. Dadas $A = \int_0^1 e^{-t^2} dt$ y $B = \int_0^{1/2} e^{-t^2} dt$. Calcular la integral reiterada.

$$I = 2\int_{-1/2}^{1} \left[\int_0^x e^{-y^2} dy \right] dx$$

en función de A y B. Existen enteros positivos m y n tales que

$$I = mA - nB + e^{-1} - e^{-1/4}.$$

Comprobar esta relación con los resultados obtenidos.

23. Un cono se obtiene uniendo todos los puntos de una región plana S con un punto no situado en el plano de S. Designando con A el área de S, y con b la altura del cono. Demostrar que:

a) El área de la sección producida por un plano paralelo a la base y a distancia t del vértice es $(t/h)^2 A$, si $0 \leq t \leq h$.

b) El volumen del cono es $\frac{1}{3} Ah$.

24. Invertir el orden de integración para deducir la fórmula

$$\int_0^a \left[\int_0^y e^{m(a-x)} f(x)\, dx \right] dy = \int_0^a (a - x) e^{m(a-x)} f(x)\, dx ,$$

donde a y m son constantes, $a > 0$.

11.16 Otras aplicaciones de las integrales dobles

Ya hemos visto que las integrales dobles pueden usarse para calcular volúmenes de sólidos y áreas de regiones planas. Otros muchos conceptos tales como, masas, centros de gravedad y momentos de inercia pueden definirse y calcularse con el auxilio de integrales dobles. En esta sección se exponen brevemente estos problemas que tienen especial importancia en física e ingeniería.

Sea $\boldsymbol{P}$ el vector que une el origen con un punto cualquiera de E_3. Si n masas positivas $m_1, m_2, \ldots, m_n$ están localizadas en los puntos $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_n$ respec-

tivamente, el *centro de gravedad* del sistema se define como el punto C determinado por el vector

$$C = \frac{\sum_{k=1}^{n} m_k \boldsymbol{P}_k}{\sum_{k=1}^{n} m_k}.$$

El denominador, $\sum m_k$, se llama *masa total* del sistema.

Si cada masa m_k se traslada según un vector A a un nuevo punto Q_k siendo $Q_k = P_k + A$, el centro de gravedad también experimenta la traslación A, puesto que se tiene

$$\frac{\sum m_k Q_k}{\sum m_k} = \frac{\sum m_k(P_k + A)}{\sum m_k} = \frac{\sum m_k P_k}{\sum m_k} + A = C + A.$$

Esto también se expresa diciendo que la posición del centro de gravedad depende tan sólo de los puntos $P_1, P_2, \ldots, P_n$ y de las masas, y no de la situación del origen. El centro de gravedad es una cantidad calculada teóricamente que representa, por así decirlo, un ficticio «punto de equilibrio» del sistema.

Si las masas están en un plano en los puntos de coordenadas $(x_1, y_1), \ldots, (x_n, y_n)$, y si el centro de gravedad tiene coordenadas $(\bar{x}, \bar{y})$, la relación vectorial que define C puede expresarse con dos ecuaciones escalares,

$$\bar{x} = \frac{\sum m_k x_k}{\sum m_k} \qquad \text{y} \qquad \bar{y} = \frac{\sum m_k y_k}{\sum m_k}.$$

En el numerador del cociente que define $\bar{x}$, el término k-ésimo de la suma, $m_k x_k$, se denomina el *momento* de la masa m_k respecto al eje y. Si una masa m igual a la masa total del sistema fuese situada en el centro de gravedad, su momento respecto al eje y sería igual al momento del sistema

$$m\bar{x} = \sum_{k=1}^{n} m_k x_k.$$

Cuando nos referimos a un sistema cuya masa total está distribuida en una cierta región del plano en lugar de estar situada en un número finito de puntos aislados, los conceptos de masa, centro de gravedad y momento se definen mediante integrales en lugar de hacerlo con sumas. Por ejemplo, consideremos una «fina lámina» que tenga la forma de una región plana S. Supongamos que la materia está distribuida por toda la lámina con una densidad conocida (masa por unidad de área). Con ello significamos que existe una función f no negativa definida en S y que $f(x, y)$ representa la masa por unidad de área en el punto (x, y). Si la lámina o placa está construida con un material homogéneo la densidad es

constante. En tal caso la masa total de la lámina se define como el producto de la densidad por el área de la lámina.

Cuando la densidad varía de un punto a otro utilizamos la integral doble de la densidad como definición de la masa total. Es decir, si la función densidad f es integrable en S, definimos la masa total $m(S)$ de la lámina mediante la fórmula

$$m(S) = \iint_S f(x, y)\, dx\, dy .$$

El cociente

$$\frac{\text{masa}}{\text{área}} = \frac{\iint_S f(x, y)\, dx\, dy}{\iint_S dx\, dy}$$

se llama *densidad media* de la lámina. Si consideramos que S no es precisamente una lámina delgada, sino una figura geométrica de dos dimensiones, el anterior cociente se llama *promedio* o *valor medio* de la función f sobre la región S. En este caso no exigimos que f sea no negativa.

Por analogía con el caso finito, definimos el *centro de gravedad* de la lámina como el punto $(\bar{x}, \bar{y})$ determinado por las fórmulas

$$(11.14)\quad \bar{x}m(S) = \iint_S xf(x, y)\, dx\, dy \qquad \text{y} \qquad \bar{y}m(S) = \iint_S yf(x, y)\, dx\, dy .$$

Las integrales del segundo miembro son los momentos de la lámina respecto a los ejes y y x respectivamente. Cuando la densidad es constante, $f(x, y) = c$, se simplifica las igualdades (11.14) y obtenemos

$$\bar{x}a(S) = \iint_S x\, dx\, dy \qquad \text{y} \qquad \bar{y}a(S) = \iint_S y\, dx\, dy ,$$

donde $a(S)$ es el área de S. En este caso el punto $(\bar{x}, \bar{y})$ se denomina el *centroide* de la lámina (o de la región S).

Si L es una recta en el plano de la lámina y representamos con $\delta(x, y)$ la distancia desde el punto (x, y) de S a la recta L, el número I_L definido por

$$I_L = \iint_S \delta^2(x, y) f(x, y)\, dx\, dy$$

se llama *momento de inercia* de la lámina respecto a L. Si $f(x, y) = 1$, I_L se denomina momento de inercia o *segundo momento* de la región S respecto a L. Los

momentos de inercia respecto a los ejes x e y se designan con I_x e I_y, respectivamente, y vienen dados por las integrales

$$I_x = \iint_S y^2 f(x, y)\, dx\, dy \qquad \text{y} \qquad I_y = \iint_S x^2 f(x, y)\, dx\, dy .$$

La suma de esas dos integrales se llama *momento polar de inercia* I_0 respecto al origen:

$$I_0 = I_x + I_y = \iint_S (x^2 + y^2) f(x, y)\, dx\, dy .$$

Nota. La masa y el centro de gravedad de una lámina tienen propiedades que dependen del cuerpo y no de la situación del origen y de las direcciones de los ejes coordenados. El momento polar de inercia depende de la posición del origen, pero no de las direcciones de los ejes. Los momentos y momentos de inercia respecto a los ejes x e y dependen de la situación del origen y de las direcciones de los ejes. Si una lámina de densidad constante posee un eje de simetría, su centroide estará en dicho eje. Si tiene dos ejes de simetría, el centroide estará en su intersección. Estas propiedades pueden demostrarse a partir de las definiciones que preceden, y son muy útiles, pues simplifican frecuentemente los cálculos.

EJEMPLO 1. Una lámina delgada de densidad constante c está limitada por dos circunferencias concéntricas de radios a y b y centro en el origen, siendo $0 < b < a$. Calcular el momento polar de inercia.

Solución. La integral para calcular I_0 es:

$$I_0 = c \iint_S (x^2 + y^2)\, dx\, dy ,$$

siendo $S = \{(x, y) \mid b^2 \leq x^2 + y^2 \leq a^2\}$. Para simplificar los cálculos tengamos en cuenta que la integral es una función aditiva en la región de integración (ya que el integrando es no negativo), de modo que tenemos:

$$I_0 = c \iint_{S(a)} (x^2 + y^2)\, dx\, dy - c \iint_{S(b)} (x^2 + y^2)\, dx\, dy ,$$

donde $S(a)$ y $S(b)$ son los discos circulares de radios a y b, respectivamente. Podemos usar la integración reiterada para calcular la integral extendida a $S(a)$, y encontramos

$$\iint_{S(a)} (x^2 + y^2)\, dx\, dy = 4 \int_0^a \left[\int_0^{\sqrt{a^2 - x^2}} (x^2 + y^2)\, dy \right] dx = \frac{\pi a^4}{2} .$$

(Omitimos los detalles de cálculo debido a que esta integral puede calcularse más fácilmente en coordenadas polares, lo que se expondrá en la sección 11.27.) Por lo tanto,

$$I_0 = \frac{\pi c}{2}(a^4 - b^4) = \pi c(a^2 - b^2)\frac{(a^2 + b^2)}{2} = m\frac{a^2 + b^2}{2},$$

donde $m = \pi c(a^2 - b^2)$ es la masa de la lámina.

EJEMPLO 2. Determinar el centroide de la región plana limitada por un arco de sinusoide.

Solución. Consideremos la región S limitada por la curva $y = \operatorname{sen} x$ y el intervalo $0 \leq x \leq \pi$. Por la simetría, la coordenada x del centroide es $\bar{x} = \pi/2$. La coordenada y viene dada por

$$\bar{y} = \frac{\iint_S y\,dx\,dy}{\iint_S dx\,dy} = \frac{\int_0^\pi [\int_0^{\operatorname{sen} x} y\,dy]\,dx}{\int_0^\pi \operatorname{sen} x\,dx} = \frac{\int_0^\pi \frac{1}{2}\operatorname{sen}^2 x\,dx}{2} = \frac{\pi}{8}.$$

11.17 Dos teoremas de Pappus

Pappus de Alejandría, que vivió en la segunda mitad del siglo tercero, fue uno de los últimos geómetras de la escuela de Alejandría de matemáticos griegos. Escribió un compendio de ocho libros en los que recogió gran parte de los conocimientos matemáticos de aquel tiempo. (Se conservan los seis últimos y una parte del segundo.) Pappus descubrió numerosas propiedades interesantes de los centroides, dos de las cuales exponemos en esta sección. La primera relaciona el centroide de una región plana con el volumen del sólido de revolución obtenido por la rotación de la citada región alrededor de una recta de su plano.

Consideremos una región plana Q situada entre las gráficas de dos funciones continuas f y g en el intervalo $[a, b]$, siendo $0 \leq g \leq f$. Sea S el sólido de revolución engendrado al girar Q alrededor del eje x. Designemos por $a(Q)$ el área de Q, $v(S)$ el volumen de S, y con $\bar{y}$ la coordenada y del centroide de Q. Al girar Q para engendrar S, el centroide se desplaza a lo largo de una circunferencia de radio $\bar{y}$. El teorema de Pappus establece que el *volumen de S es igual al producto de la longitud de esa circunfarencia multiplicada por el área de Q*; esto es,

(11.15) $$v(S) = 2\pi\bar{y}a(Q).$$

Para demostrar esta fórmula observemos tan sólo que el volumen viene dado por la integral

$$v(S) = \pi \int_a^b [f^2(x) - g^2(x)]\, dx$$

y que $\bar{y}$ se obtiene con la fórmula

$$\bar{y}a(Q) = \iint_Q y\, dy\, dx = \int_a^b \left[\int_{g(x)}^{f(x)} y\, dy \right] dx = \int_a^b \tfrac{1}{2}[f^2(x) - g^2(x)]\, dx\,.$$

Comparando esas dos fórmulas se obtiene inmediatamente (11.15).

EJEMPLO 1. *Volumen de un toro.* Sea S el toro engendrado al girar un disco circular Q de radio R alrededor de un eje a una distancia $b > R$ del centro de Q. El volumen de S se calcula fácilmente con el teorema de Pappus. Tenemos $\bar{y} = b$ y $a(Q) = \pi R^2$, así que

$$v(S) = 2\pi \bar{y} a(Q) = 2\pi^2 R^2 b\,.$$

El ejemplo que sigue hace ver que el teorema de Pappus también puede usarse para determinar centroides.

EJEMPLO 2. *Centroide de un disco semicircular.* Sea $\bar{y}$ la coordenada y del centroide del disco semicircular

$$Q = \{(x, y) \mid x^2 + y^2 \le R^2, \quad y \ge 0\}\,.$$

El área de Q es $\frac{1}{2}\pi R^2$. Cuando Q gira alrededor del eje x engendra una esfera de volumen $\frac{4}{3}\pi R^3$. Según la fórmula de Pappus tenemos

$$\tfrac{4}{3}\pi R^3 = 2\pi \bar{y}(\tfrac{1}{2}\pi R^2)\,,$$

así que $\bar{y} = \dfrac{4R}{3\pi}\cdot$

El teorema de Pappus que sigue establece que el *centroide de la reunión de dos regiones planas disjuntas A y B está en el segmento de recta que une el centroide de A con el centroide de B.* Con mayor generalidad, sean A y B dos láminas delgadas que son disjuntas o se cortan en un conjunto de puntos de contenido nulo. Designemos con $m(A)$ y $m(B)$ sus masas y con C_A y C_B los vectores que unen el origen con los respectivos centros de gravedad. La reunión $A \cup B$

tiene entonces masa $m(A) + m(B)$ y su centro de gravedad está determinado por el vector C dado por la fórmula

(11.16) $$C = \frac{m(A)C_A + m(B)C_B}{m(A) + m(B)}.$$

El cociente que da C es una combinación lineal de la forma $aC_A + bC_B$, en la que a y b son escalares no negativos cuya suma es 1. Una combinación lineal de esta forma se llama *combinación convexa* de C_A y C_B. El extremo de C está situado en el segmento de recta que une los extremos de C_A y C_B.

La fórmula de Pappus (11.16) se deduce inmediatamente de la definición de centro de gravedad que se da en (11.14). Puede generalizarse de manera sencilla a la reunión de tres o más regiones. Es útil especialmente en la práctica cuando una lámina de densidad constante se compone de varias porciones, cada una de las cuales tiene simetría geométrica. Determinamos el centroide de cada porción y formamos luego una adecuada combinación lineal para encontrar el centroide de la reunión. En el ejercicio 21 de la sección siguiente se dan ejemplos.

11.18 Ejercicios

En los ejercicios del 1 al 8 la región S está limitada por una o más curvas de las que se dan las ecuaciones. En cada caso dibujar la región S y determinar las coordenadas $\bar{x}$ e $\bar{y}$ del centroide.

1. $y = x^2, \quad x + y = 2.$
2. $y^2 = x + 3, \quad y^2 = 5 - x.$
3. $x - 2y + 8 = 0, \quad x + 3y + 5 = 0, \quad x = -2, \quad x = 4.$
4. $y = \operatorname{sen}^2 x, \quad y = 0, \quad 0 \leq x \leq \pi.$
5. $y = \operatorname{sen} x, \quad y = \cos x, \quad 0 \leq x \leq \frac{\pi}{4}.$
6. $y = \log x, \quad y = 0, \quad 1 \leq x \leq a.$
7. $\sqrt{x} + \sqrt{y} = 1, \quad x = 0, \quad y = 0.$
8. $x^{2/3} + y^{2/3} = 1, \quad x = 0, \quad y = 0,$ en el primer cuadrante.
9. Una lámina delgada está limitada por el arco de parábola $y = 2x - x^2$ y el intervalo $0 \leq x \leq 2$. Determinar su masa si la densidad en cada punto (x, y) es $(1 - y)/(1 + x)$.
10. Determinar el centro de gravedad de una lámina delgada rectangular $ABCD$ si la densidad en todos sus puntos es el producto de sus distancias a los lados AB y AD.

En los ejercicios 11 a 16 calcular los momentos de inercia I_x e I_y de una lámina delgada S del plano xy bordeada por una o más curvas cuyas ecuaciones se citan. En cada caso $f(x, y)$ representa la densidad de un punto cualquiera (x, y) de S.

11. $y = \operatorname{sen}^2 x, \quad y = -\operatorname{sen}^2 x, \quad -\pi \leq x \leq \pi; \quad f(x, y) = 1.$

12. $\frac{x}{a}+\frac{y}{b}=1, \quad \frac{x}{c}+\frac{y}{b}=1, \quad y=0, \quad 0<c<a, \quad b>0; \quad f(x,y)=1.$

13. $(x-r)^2+(y-r)^2=r^2, \quad x=0, \quad y=0, \quad 0\le x\le r, 0\le y\le r; f(x,y)=1.$

14. $xy=1, \quad xy=2, \quad x=2y, \quad y=2x, \quad x>0, \quad y>0; \quad f(x,y)=1.$

15. $y=e^x, \quad y=0, \quad 0\le x\le a; \quad f(x,y)=xy.$

16. $y=\sqrt{2x}, \quad y=0, \quad 0\le x\le 2; \quad f(x,y)=|x-y|.$

17. Sean S una lámina delgada de masa m, y L_0 y L dos rectas paralelas en el plano de S, pasando L_0 por el centro de gravedad de S. Demostrar el *teorema de los ejes paralelos:*

$$I_L = I_{L_0} + mh^2,$$

donde h es la distancia entre las dos rectas L_0 y L. [*Indicación.* El trabajo se simplificará con una adecuada elección de los ejes coordenados.]

18. El contorno de una lámina delgada es una elipse de semiejes a y b. L representa una recta en el plano de la lámina que pasa por el centro de la elipse y forma un ángulo α con el eje de longitud $2a$. Si la densidad es constante y la masa m, demostrar que el momento de inercia I_L es igual a $\frac{1}{4}m(a^2\operatorname{sen}^2\alpha + b^2\cos^2\alpha)$.

19. Encontrar la distancia media desde un vértice de un cuadrado de lado h a los puntos interiores del mismo.

20. Sea δ la distancia desde un punto arbitrario P interior a un círculo de radio r a un punto fijo P_0 cuya distancia al centro del círculo es h. Calcular el valor medio de la función δ^2 en la región limitada por el círculo.

21. Sean A, B, C los siguientes rectángulos del plano xy:

$$A=[0,4]\times[0,1], \qquad B=[2,3]\times[1,3], \qquad C=[2,4]\times[3,4].$$

Aplicar el teorema de Pappus para determinar el centroide de las siguientes figuras:

a) $A\cup B$.
b) $A\cup C$.
c) $B\cup C$.
d) $A\cup B\cup C$.

22. Un triángulo isósceles T tiene base 1 y altura h. La base de T coincide con uno de los lados de un rectángulo R de base 1 y altura 2. Encontrar el valor de h de manera que el centroide de $R\cup T$ esté situado en el lado común de R y T.

23. Un triángulo isósceles T tiene base $2r$ y altura r. La base de T coincide con el lado recto de un disco semicircular D de radio r. Determinar la relación que debe existir entre r y h para que el centroide de $T\cup D$ esté situado en el interior del triángulo.

11.19 Teorema de Green en el plano

El segundo teorema fundamental del cálculo para las integrales de línea establece que la integral de línea de un gradiente ∇f a lo largo de un camino que une dos puntos $\boldsymbol{a}$ y $\boldsymbol{b}$ puede expresarse en función de los valores $f(\boldsymbol{a})$ y $f(\boldsymbol{b})$. Existe un teorema análogo en dos dimensiones que expresa una integral doble extendida a una región R como una integral de línea a lo largo de la curva cerrada que constituye la frontera de R. Este teorema corrientemente se denomina *teorema*

de Green, (*) pero a veces se llama teorema de Gauss. Puede establecerse de varias maneras; la más frecuente es en la forma de la identidad:

$$\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \oint_C P\, dx + Q\, dy. \tag{11.17}$$

La curva C que aparece en el segundo miembro es el contorno de la región R, y el símbolo de integración $\oint$ indica que la curva se recorre en el sentido contrario al de las agujas del reloj, como se ve en la figura 11.13.

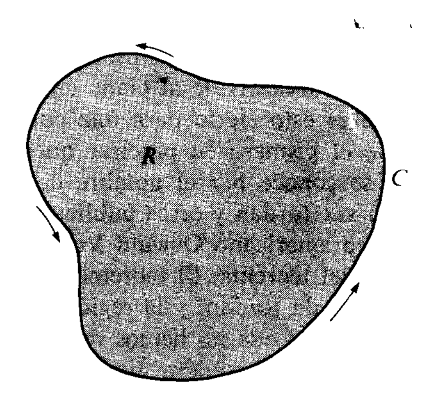


FIGURA 11.13 *La curva C es la frontera de R, recorrida en el sentido contrario al de las agujas del reloj.*

Se precisan dos tipos de hipótesis para la validez de la identidad. Primero, se imponen condiciones a P y Q a fin de asegurar la existencia de las integrales. Las corrientes son que P y Q sean derivables con continuidad en un conjunto abierto S que contenga la región R. Esto implica la continuidad de P y Q sobre C así como la continuidad de $\partial P/\partial y$ y $\partial Q/\partial x$ en R, aunque el teorema también es cierto con hipótesis menos restrictivas. Segundo, hay otras condiciones de tipo geométrico que se imponen a la región R y a su contorno C. La curva C puede ser cualquier *curva cerrada simple rectificable.* La palabra «rectificable» significa, naturalmente, que C tiene una longitud finita. A fin de explicar lo que se entiende por curva cerrada simple, nos referiremos a la función vectorial que describe la curva.

(*) En honor a George Green (1793-1841), matemático inglés que estudió las aplicaciones matemáticas a la electricidad y magnetismo, flujo de fluidos, y a la reflexión y refracción de la luz y del sonido. El teorema que lleva el nombre de Green aparece ya en las investigaciones de Gauss y de Lagrange.

Supongamos que C viene descrita por una función vectorial continua $\boldsymbol{\alpha}$ definida en un intervalo $[a,b]$. Si $\boldsymbol{\alpha}(a) = \boldsymbol{\alpha}(b)$, la curva es *cerrada.* Una curva cerrada tal que $\boldsymbol{\alpha}(t_1) \neq \boldsymbol{\alpha}(t_2)$ para todo par de valores $t_1 \neq t_2$ del intervalo semiabierto $(a, b]$ se llama curva cerrada *simple.* Esto significa que, excepto para los extremos del intervalo $[a, b]$, valores distintos de t originan puntos distintos de la curva. La circunferencia es el típico ejemplo de curva cerrada simple.

Las curvas cerradas simples planas se llaman corrientemente *curvas de Jordan* en recuerdo de Camilo Jordan (1838-1922), famoso matemático francés que fue de los primeros que trabajó sobre los conceptos de curvas cerradas simples y de longitud de arco. Toda curva de Jordan C descompone el plano en dos conjuntos abiertos conexos y disjuntos que tienen la curva C como frontera común. Una de esas regiones es *acotada* y se llama *interior* a C. (Véase la región sombreada en la figura 11.13.) La otra es *no acotada* y se llama *exterior* a C. Para ciertas curvas de Jordan familiares como circunferencias, elipses, o polígonos elementales, es intuitivamente evidente que la curva divide al plano en una región interior y otra exterior, pero demostrar que es esto cierto para una curva de Jordan *cualquiera* no es cosa fácil. Jordan fue el primero en precisar que esa afirmación requería demostración; el resultado se conoce por el nombre de *teorema de la curva de Jordan.* Hacia fines del siglo XIX Jordan y otros publicaron demostraciones incompletas. En 1905 el matemático americano Oswald Veblen (1880-1960) dio la primera demostración completa del teorema. El teorema de Green es válido siempre que C es una curva rectificable de Jordan, y la región R es la reunión de C y de la región interior a ella. (*) Puesto que no hemos definido las integrales de línea a lo largo de curvas rectificables cualesquiera, limitamos nuestra discusión a curvas regulares a trozos.

Existe otra dificultad técnica asociada a la formulación del teorema de Green. Ya hemos observado que, para la validez de la identidad (11.17), la curva C debe ser recorrida en sentido contrario al de las agujas del reloj. Intuitivamente, esto significa que un hombre que vaya andando a lo largo de C tiene siempre la región R a su izquierda. También, para algunas curvas de Jordan corrientes, como las antes mencionadas, el significado de la expresión «recorrer una curva en sentido contrario al de las agujas del reloj» es intuitivamente evidente. No obstante, en una demostración estrictamente rigurosa del teorema de Green se debería definir esta expresión en términos completamente analíticos, esto es, mediante la función vectorial $\boldsymbol{\alpha}$ que describe la curva. En la sección 11.24 se esboza una posible definición.

Habiendo ya señalado algunas de las dificultades asociadas a la formulación del teorema de Green, lo enunciaremos en una forma más bien general e indicaremos brevemente por qué es válido para ciertas regiones. En esta discusión el

(*) En el capítulo 10 de la obra del autor *Análisis matemático,* se puede ver una demostración del teorema de Green para regiones de tal generalidad.

significado de «contrario al de las agujas del reloj» es intuitiva, de manera que la demostración no es completamente rigurosa.

TEOREMA 11.10. TEOREMA DE GREEN PARA REGIONES PLANAS LIMITADAS POR CURVAS DE JORDAN REGULARES A TROZOS. *Sean P y Q campos escalares derivables con continuidad en un conjunto abierto S del plano xy. Sea C una curva de Jordan regular a trozos, y representemos por R la reunión de C y de su interior. Supongamos que R está contenida en S. Se tiene entonces la identidad*

$$\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \oint_C P\, dx + Q\, dy, \tag{11.18}$$

en la que la integral de línea se toma alrededor de C en sentido contrario al de las agujas del reloj.

Nota. La identidad (11.18) es equivalente a las *dos* fórmulas

$$\iint_R \frac{\partial Q}{\partial x}\, dx\, dy = \oint_C Q\, dy \tag{11.19}$$

y

$$-\iint_R \frac{\partial P}{\partial y}\, dx\, dy = \oint_C P\, dx. \tag{11.20}$$

En efecto, si estas dos fórmulas son válidas, sumando se deduce (11.18). Recíprocamente, si (11.18) es cierta podemos obtener (11.19) y (11.20) como caso particular tomando $P = 0$ y $Q = 0$, respectivamente.

Demostración para regiones especiales. Demostraremos (11.20) para una región R del tipo I. Una tal región tiene la forma

$$R = \{(x, y) \mid a \leq x \leq b \quad \text{y} \quad f(x) \leq y \leq g(x)\},$$

en donde f y g son continuas en $[a, b]$ siendo $f \leq g$. La frontera C de R consta de cuatro partes, un arco inferior C_1 (gráfica de f), otro superior C_2 (gráfica de g), y dos segmentos rectilíneos verticales, recorridos en el sentido indicado en la figura 11.14.

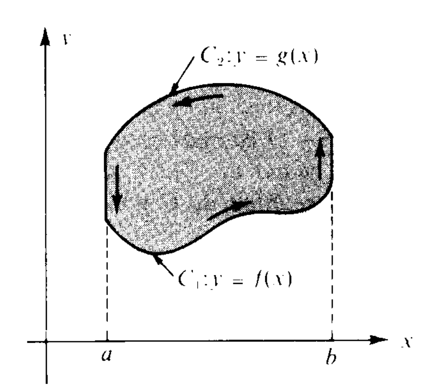


FIGURA 11.14 *Demostración del teorema de Green para una región especial.*

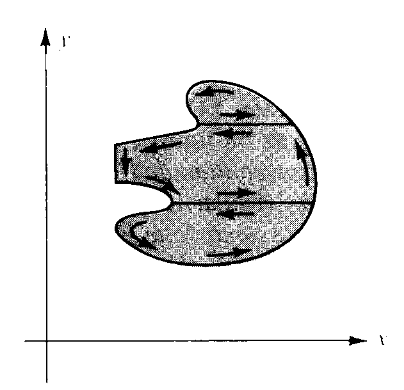


FIGURA 11.15 *Demostración del teorema de Green para una región más general.*

Calculemos primero la integral doble $-\iint_R (\partial P/\partial y)\, dx\, dy$ por integración sucesiva. Integrando primero respecto a y tenemos

$$(11.21) \qquad -\iint_R \frac{\partial P}{\partial y}\, dx\, dy = -\int_a^b \left[\int_{f(x)}^{g(x)} \frac{\partial P}{\partial y}\, dy\right] dx = \int_a^b \left[\int_{g(x)}^{f(x)} \frac{\partial P}{\partial y}\, dy\right] dx$$

$$= \int_a^b P[x, f(x)]\, dx - \int_a^b P[x, g(x)]\, dx\,.$$

Por otra parte, la integral de línea $\int_C P\, dx$ puede escribirse así

$$\int_C P\, dx = \int_{C_1} P\, dx + \int_{C_2} P\, dx\,,$$

puesto que la integral de línea a lo largo de cada segmento vertical es cero. Para calcular la integral a lo largo de C_1 utilizamos la representación vectorial $\alpha(t) = ti + f(t)j$ y obtenemos

$$\int_{C_1} P\, dx = \int_a^b P[t, f(t)]\, dt\,.$$

Para calcular la integral a lo largo de C_2 emplearemos la representación $\alpha(t) = ti + g(t)j$ y tendremos en cuenta la inversión del sentido para obtener

$$\int_{C_2} P\, dx = -\int_a^b P[t, g(t)]\, dt\,.$$

Tenemos por lo tanto

$$\int_C P\, dx = \int_a^b P[t, f(t)]\, dt - \int_a^b P[t, g(t)]\, dt\,.$$

Comparando esta fórmula con la (11.21) obtenemos (11.20).

Un razonamiento parecido puede emplearse para demostrar (11.19) para regiones del tipo II. Se obtiene de este modo una demostración del teorema de Green para regiones que son a un tiempo del tipo I y del tipo II. Hecho esto, puede demostrarse el teorema para regiones R que pueden descomponerse en un número finito de regiones que son de ambos tipos. Se introducen «secciones» o «cortes» como se ve en la figura 11.15, se aplica el teorema a cada subregión, y se suman los resultados. Las integrales a lo largo de los cortes se reducen a pares, como se sugiere en la figura, y la suma de las integrales de línea a lo largo de las fronteras de las subregiones es igual a la integral de línea a lo largo de la frontera de R.

11.20 Algunas aplicaciones del teorema de Green

Los ejemplos que siguen muestran algunas aplicaciones del teorema de Green.

EJEMPLO 1. Por medio del teorema de Green calcular el trabajo efectuado por el campo de fuerza $f(x, y) = (y + 3x)i + (2y - x)j$ al mover una partícula rodeando una vez la elipse $4x^2 + y^2 = 4$ en el sentido contrario al de las agujas del reloj.

Solución. El trabajo es igual $\int_C P\,dx + Q\,dy$, donde $P = y + 3x$, $Q = 2y - x$, y C es la elipse. Ya que $\partial Q/\partial x - \partial P/\partial y = -2$, el teorema de Green nos da

$$\int_C P\,dx + Q\,dy = \iint_R (-2)\,dx\,dy = -2a(R),$$

donde $a(R)$ es el área interior a la elipse. Ya que esta elipse tiene semiejes $a = 1$ y $b = 2$, su área es $\pi ab = 2\pi$ y el valor de la integral de línea es -4π.

EJEMPLO 2. Calcular la integral $\int_C (5 - xy - y^2)dx - (2xy - x^2)dy$, donde C es el cuadrado de vértices (0, 0), (1, 0), (1, 1), (0, 1).

Solución. Aquí $P = 5 - xy - y^2$, $Q = x^2 - 2xy$, y $\partial Q/\partial x - \partial P/\partial y = 3x$. Luego, según el teorema de Green, tenemos,

$$\int_C P\,dx + Q\,dy = 3\iint_R x\,dx\,dy = 3\bar{x},$$

donde $\bar{x}$ es la coordenada x del centroide del cuadrado. Puesto que $\bar{x} = \frac{1}{2}$, el valor de la integral de línea es $\frac{3}{2}$.

EJEMPLO 3. *El área expresada como una integral de línea.* La integral do-

ble que da el área $a(R)$ de una región R puede expresarse en la forma

$$a(R) = \iint_R dx\, dy = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy\,,$$

en donde P y Q son tales que $\partial Q/\partial x - \partial P/\partial y = 1$. Por ejemplo, podemos tomar $Q(x, y) = \frac{1}{2}x$ y $P(x, y) = -\frac{1}{2}y$. Si R es la región encerrada en una curva de Jordan C podemos aplicar el teorema de Green para expresar $a(R)$ como integral de línea

$$a(R) = \int_C P\, dx + Q\, dy = \frac{1}{2}\int_C -y\, dx + x\, dy\,.$$

Si la curva frontera C está dada paramétricamente

$$x = X(t), \qquad y = Y(t), \qquad a \le t \le b\,,$$

la integral de línea que da el área toma la forma

$$a(R) = \frac{1}{2}\int_a^b \{-Y(t)X'(t) + X(t)Y'(t)\}\, dt = \frac{1}{2}\int_a^b \begin{vmatrix} X(t) & Y(t) \\ X'(t) & Y'(t) \end{vmatrix} dt\,.$$

11.21 Condición necesaria y suficiente para que un campo vectorial bi-dimensional sea un gradiente

Sea $f(x, y) = P(x, y)i + Q(x, y)j$ un campo vectorial derivable con continuidad en un conjunto abierto S del plano. Si f es un gradiente en S tenemos

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{11.22}$$

en todo S. Es decir, la condición (11.22) es *necesaria* para que f sea un gradiente. Como ya hemos observado, esta condición no es suficiente. Por ejemplo, el campo vectorial

$$f(x, y) = \frac{-y}{x^2 + y^2}\,i + \frac{x}{x^2 + y^2}\,j$$

satisface (11.22) en todo el conjunto $S = \mathbf{R}^2 - \{(0, 0)\}$, pero f no es un gradiente

en S. En el teorema 10.9 se demostró que la condición (11.22) es a la vez necesaria y suficiente para que f sea un gradiente en S si el conjunto S es *convexo.* Con la ayuda del teorema de Green podemos extender este resultado a una clase más general de conjuntos planos como los conjuntos *simplemente conexos.* Éstos se definen como sigue.

DEFINICIÓN DE CONJUNTO PLANO SIMPLEMENTE CONEXO. *Sea S un conjunto plano abierto y conexo. Se dice que S es simplemente conexo si, para toda curva de Jordan C situada en S, la región interior a S es también un subconjunto de S.*

Un anillo (conjunto de puntos situados entre dos circunferencias concéntricas) no es simplemente conexo debido a que toda la región interior a la circunferencia externa no es un subconjunto del anillo. Dicho en forma intuitiva, un conjunto S es simplemente conexo cuando no tiene «agujeros». Otro modo de describir la conexión simple consiste en decir que una curva C_1 de S que une dos puntos cualesquiera puede ser transformada sin ruptura en otra curva C_2 de S que une aquellos dos puntos, con la particularidad de que todas las curvas intermedias obtenidas en la deformación están contenidas en S. Otra definición que es equivalente a la dada, según puede demostrarse, dice que un conjunto es simplemente conexo si su complementario (respecto a todo el plano) no es conexo. Por ejemplo, un anillo no es simplemente conexo porque su complemento es no conexo. Un conjunto abierto y conexo que no es simplemente conexo se llama múltiplemente conexo.

TEOREMA 11.11. *Si $f(x, y)=P(x, y)\boldsymbol{i}+Q(x, y)\boldsymbol{j}$ es un campo vectorial derivable con continuidad en un conjunto abierto simplemente conexo S del plano, entonces f es un gradiente en S si y sólo si tenemos*

$$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x} \quad \textit{en todos los puntos de } S. \tag{11.23}$$

Demostración. Ya hemos dicho que la condición (11.23) es necesaria para que f sea un gradiente. Demostraremos ahora que también es suficiente.

Puede demostrarse que en cualquier conjunto plano conexo y abierto S, todo par de puntos $\boldsymbol{a}$ y $\boldsymbol{x}$ puede unirse mediante un polígono escalonado simple, esto es, mediante un polígono cuyos lados son paralelos a los ejes coordenados y que no se cortan a sí mismos. Si la integral de línea de f desde $\boldsymbol{a}$ hasta $\boldsymbol{x}$ tiene el mismo valor para todo polígono escalonado simple en S que una $\boldsymbol{a}$ con $\boldsymbol{x}$, entonces el mismo razonamiento empleado para demostrar el teorema 10.4 prueba que f es un gradiente en S. Por consiguiente, necesitamos tan sólo comprobar que la integral de línea de f desde $\boldsymbol{a}$ hasta $\boldsymbol{x}$ tiene el mismo valor para todo polígono escalonado simple en S que una $\boldsymbol{a}$ con $\boldsymbol{x}$.

Sean C_1 y C_2 dos polígonos escalonados simples en S que unan $\boldsymbol{a}$ y $\boldsymbol{x}$. Algu-

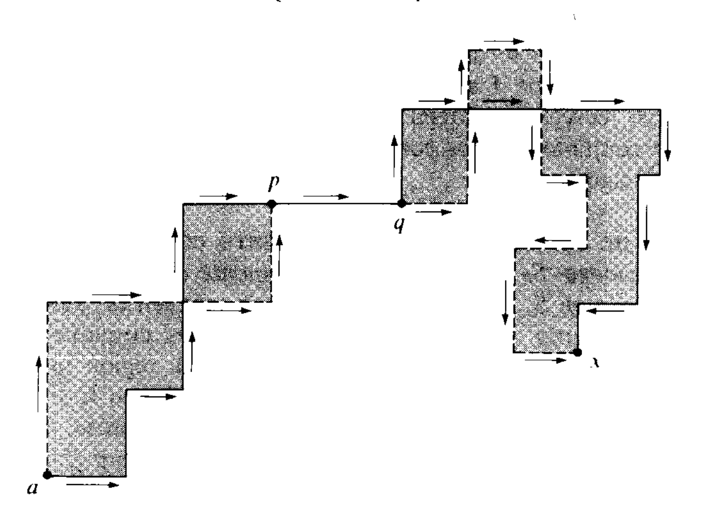


FIGURA 11.16 *Independencia del camino en una región simplemente conexa*

nas porciones de estos polígonos pueden coincidir. Las restantes porciones se cortarán a lo sumo un número finito de veces, y formarán las fronteras de un número finito de regiones poligonales, $R_1, \ldots, R_m$. Puesto que se supone que S es simplemente conexo, cada una de las regiones R_k es un subconjunto de S. En la figura 11.16 se ha dibujado un ejemplo. La poligonal de trazo continuo representa C_1 y la de puntos C_2, las regiones sombreadas son las $R_1, \ldots, R_m$. (A lo largo del segmento $\boldsymbol{pq}$ esas dos poligonales coinciden.)

Observemos ahora que la integral de línea de $\boldsymbol{f}$ a lo largo de C_1 desde $\boldsymbol{a}$ hasta $\boldsymbol{x}$ más la integral de $\boldsymbol{x}$ hasta $\boldsymbol{a}$ a lo largo de C_2 es cero ya que la integral a lo largo del camino cerrado es una suma de integrales tomadas sobre los segmentos comunes a C_1 y C_2 más las integrales tomadas a lo largo de las fronteras de las regiones R_k. Las calculadas sobre los segmentos comunes se anulan a pares, ya que cada uno de tales segmentos se recorre dos veces, en sentidos opuestos, y su suma es cero. La integral sobre la frontera Γ_k de cada región R_k también es cero porque, en virtud del teorema de Green, podemos escribir

$$\int_{\Gamma_k} P\,dx + Q\,dy = \pm \iint_{R_k} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy\,,$$

y el integrando de la integral doble es cero en virtud de la hipótesis $\partial Q/\partial x = \partial P/\partial y$ Se deduce que la integral de $\boldsymbol{a}$ a $\boldsymbol{x}$ a lo largo de C_1 es igual a la calculada a lo largo de C_2. Como hemos observado antes, esto implica que $\boldsymbol{f}$ es un gradiente en S.

11.22 Ejercicios

1. Usar el teorema de Green para calcular la integral $\oint_C y^2\,dx + x\,dy$ cuando
   a) C es el cuadrado de vértices $(0,0)$, $(2,0)$, $(2,2)$, $(0,2)$.
   b) C es el cuadrado de vértices $(\pm 1, \pm 1)$.
   c) C es el cuadrado de vértices $(\pm 2, 0)$, $(0, \pm 2)$.
   d) C es la circunferencia de radio 2 y centro en el origen.
   e) C tiene la ecuación vectorial $\boldsymbol{\alpha}(t) = 2\cos^3 t\,\boldsymbol{i} + 2\operatorname{sen}^3 t\,\boldsymbol{j}$, $0 \le t \le 2\pi$.
2. Si $P(x,y) = xe^{-y^2}$ y $Q(x,y) = -x^2 y\,e^{-y^2} + 1/(x^2+y^2)$ calcular la integral $\oint P\,dx + Q\,dy$ siguiendo el contorno del cuadrado de lado $2a$ determinado por las desigualdades $|x| \le a$ y $|y| \le a$.
3. Sea C una curva cerrada simple del plano xy y representemos con I_z el momento de inercia (alrededor del eje z) de la región interior a C. Demostrar que existe un entero n tal que

$$nI_z = \oint_C x^3\,dy - y^3\,dx.$$

4. Dados dos campos escalares u y v derivables con continuidad en un conjunto abierto que contiene el disco R cuya frontera es la circunferencia $x^2 + y^2 = 1$. Definimos dos campos vectoriales $\boldsymbol{f}$ y $\boldsymbol{g}$ como sigue:

$$\boldsymbol{f}(x,y) = v(x,y)\boldsymbol{i} + u(x,y)\boldsymbol{j}, \qquad \boldsymbol{g}(x,y) = \left(\frac{\partial u}{\partial x} - \frac{\partial u}{\partial y}\right)\boldsymbol{i} + \left(\frac{\partial v}{\partial x} - \frac{\partial v}{\partial y}\right)\boldsymbol{j}.$$

   Encontrar el valor de la integral doble $\iint_R \boldsymbol{f}\cdot\boldsymbol{g}\,dx\,dy$ si se sabe que sobre la frontera de R se tiene $u(x,y) = 1$ y $v(x,y) = y$.
5. Si f y g son derivables con continuidad en un conjunto abierto conexo S del plano, demostrar que $\oint_C f\nabla g\cdot d\boldsymbol{\alpha} = -\oint_C g\nabla f\cdot d\boldsymbol{\alpha}$ para toda curva de Jordan C regular a trozos contenida en S.
6. Sean u y v dos campos escalares con derivadas parciales primeras y segundas continuas en un conjunto abierto conexo S del plano. Sea R una región de S limitada por una curva de Jordan C regular a trozos. Demostrar que

$$\text{a)}\quad \oint_C uv\,dx + uv\,dy = \iint_R \left\{v\left(\frac{\partial u}{\partial x} - \frac{\partial u}{\partial y}\right) + u\left(\frac{\partial v}{\partial x} - \frac{\partial v}{\partial y}\right)\right\} dx\,dy.$$

$$\text{b)}\quad \frac{1}{2}\oint_C \left(v\frac{\partial u}{\partial x} - u\frac{\partial v}{\partial x}\right) dx + \left(u\frac{\partial v}{\partial y} - v\frac{\partial u}{\partial y}\right) dy = \iint_R \left(u\frac{\partial^2 v}{\partial x\,\partial y} - v\frac{\partial^2 u}{\partial x\,\partial y}\right) dx\,dy.$$

Derivadas normales. En la sección 10.7 definimos las integrales de línea respecto a la longitud del arco de tal modo que es válida la siguiente igualdad:

$$\int_C P\,dx + Q\,dy = \int_C \boldsymbol{f}\cdot\boldsymbol{T}\,ds,$$

donde $\boldsymbol{f} = P\boldsymbol{i} + Q\boldsymbol{j}$ y $\boldsymbol{T}$ es el vector unitario tangente a C. (El producto interior $\boldsymbol{f}\cdot\boldsymbol{T}$ se llama componente tangencial de $\boldsymbol{f}$ a lo largo de C.) Si C es una curva de Jordan representada por una función derivable con continuidad $\boldsymbol{\alpha}(t) = X(t)\boldsymbol{i} + Y(t)\boldsymbol{j}$, el vector *normal exterior* unitario $\boldsymbol{n}$ a C se define por

$$\boldsymbol{n}(t) = \frac{1}{\|\boldsymbol{\alpha}'(t)\|}(Y'(t)\boldsymbol{i} - X'(t)\boldsymbol{j})$$

donde $\|\boldsymbol{\alpha}'(t)\| \neq 0$. Si φ es un campo escalar con un gradiente $\nabla\varphi$ en C, la *derivada normal* $\partial\varphi/\partial n$ se define en C mediante la ecuación

$$\frac{\partial\varphi}{\partial n} = \nabla\varphi\cdot\boldsymbol{n}.$$

Esto es, naturalmente, la derivada direccional de φ en la dirección de $\boldsymbol{n}$. En los ejercicios restantes de esta sección se tratan estos conceptos.

7. Si $\boldsymbol{f} = Q\boldsymbol{i} - P\boldsymbol{j}$, demostrar que

$$\int_C P\,dx + Q\,dy = \int_C \boldsymbol{f}\cdot\boldsymbol{n}\,ds.$$

(El producto $\boldsymbol{f}\cdot\boldsymbol{n}$ se llama componente normal de $\boldsymbol{f}$ a lo largo de C.)

8. Sean f y g dos campos escalares con derivadas parciales segundas y primeras continuas en un conjunto abierto S del plano. Representemos con R una región (en S) cuya frontera es una curva de Jordan C regular a trozos. Demostrar las identidades siguientes, donde $\nabla^2 u = \partial^2 u/\partial x^2 + \partial^2 u/\partial y^2$.

a) $\displaystyle\oint_C \frac{\partial g}{\partial n}\,ds = \iint_R \nabla^2 g\,dx\,dy.$

b) $\displaystyle\oint_C f\frac{\partial g}{\partial n}\,ds = \iint_R (f\nabla^2 g + \nabla f\cdot\nabla g)\,dx\,dy.$

c) $\displaystyle\oint_C \left(f\frac{\partial g}{\partial n} - g\frac{\partial f}{\partial n}\right)ds = \iint_R (f\nabla^2 g - g\nabla^2 f)\,dx\,dy.$

La identidad c) se llama *fórmula de Green;* ella demuestra que

$$\oint_C f\frac{\partial g}{\partial n}\,ds = \oint_C g\frac{\partial f}{\partial n}\,ds$$

donde f y g son armónicas en R (esto es, cuando $\nabla^2 f = \nabla^2 g = 0$ en R).

9. Supongamos que la ecuación diferencial

$$P(x,y)\,dx + Q(x,y)\,dy = 0$$

tiene un factor integrante $\mu(x,y)$ que nos permite obtener una familia simplemente infinita

de soluciones de la forma $\varphi(x, y) = C$. Si la pendiente de la curva $\varphi(x, y) = C$ en (x, y) es $\tan \theta$, el vector unitario normal $\boldsymbol{n}$ se expresa

$$\boldsymbol{n} = \operatorname{sen} \theta\, \boldsymbol{i} - \cos \theta\, \boldsymbol{j}.$$

Existe un campo escalar $g(x, y)$ tal que la derivada normal de φ viene dada por la fórmula

$$\frac{\partial \varphi}{\partial n} = \mu(x, y) g(x, y),$$

donde $\partial\varphi/\partial n = \nabla \varphi \cdot \boldsymbol{n}$. Hallar una fórmula explícita para $g(x, y)$ en función de $P(x, y)$ y $Q(x, y)$.

* 11.23 Teorema de Green para regiones múltiplemente conexas

El teorema de Green puede generalizarse y ser aplicado a ciertas regiones múltiplemente conexas.

TEOREMA 11.12. TEOREMA DE GREEN PARA REGIONES MÚLTIPLEMENTE CONEXAS. *Sean $C_1, \ldots, C_n$ n curvas de Jordan regulares a trozos que tienen las propiedades siguientes:*

a) *Dos cualesquiera de esas curvas no se cortan.*

b) *Todas las curvas $C_2, \ldots, C_n$ están situadas en el interior de C_1.*

c) *La curva C_i está en el interior de la curva C_j para cada $i \neq j$, $i > 1$, $j > 1$.*

Designemos con R la región que consiste en la reunión de C_1 con la porción del interior de C_1 que no está dentro de cualquiera de las curvas $C_2, C_3, \ldots, C_n$. (En la figura 11.17 está representada una de tales regiones.) Sean P y Q derivables con continuidad en un conjunto abierto S que contiene R. Tenemos entonces la siguiente identidad:

$$\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \oint_{C_1} (P\, dx + Q\, dy) - \sum_{k=2}^{n} \oint_{C_k} (P\, dx + Q\, dy). \tag{11.24}$$

El teorema puede demostrarse introduciendo cortes que transformen R en una reunión de un número finito de regiones simplemente conexas bordeadas por curvas de Jordan. El teorema de Green se aplica separadamente a cada una de las partes, y se suman luego los resultados. Comentaremos el caso $n = 2$. Por inducción se demuestra para un número n cualquiera de curvas.

La idea de la demostración cuando $n = 2$ se ilustra con el ejemplo de la figura 11.18, donde C_1 y C_2 son dos circunferencias, siendo C_1 la mayor. Practicamos los cortes AB y CD, como se ve en la figura. Sea K_1 la curva de Jordan consistente en

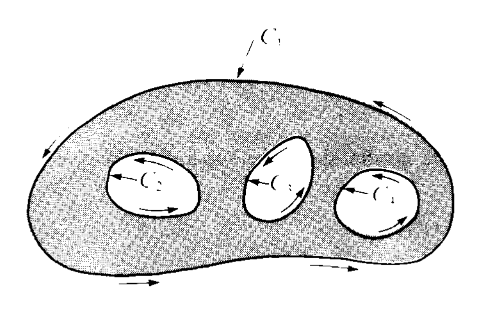


FIGURA 11.17 *Región múltiplemente conexa.*

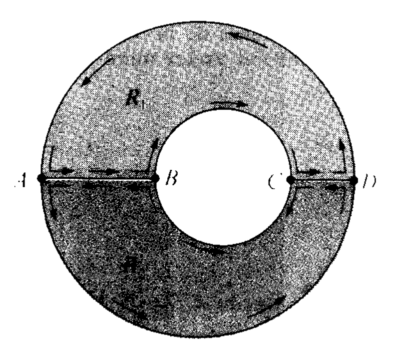


FIGURA 11.18 *Demostración del teorema de Green para una región múltiplemente conexa.*

las mitades superiores de C_1 y C_2, y los dos segmentos AB y CD. Llamemos K_2 otra curva de Jordan formada por las mitades inferiores de C_1 y C_2, y los dos segmentos AB y CD. Aplicamos el teorema de Green a cada una de las regiones limitadas por K_1 y K_2 y sumamos las dos identidades así obtenidas. En la suma las integrales calculadas a lo largo de los cortes se anulan unas con otras (pues cada corte es recorrido una vez en cada dirección), y se obtiene

$$\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \oint_{C_1} (P\, dx + Q\, dy) - \oint_{C_2} (P\, dx + Q\, dy).$$

El signo menos aparece debido a la dirección con que se recorre C_2. Esta es la ecuación (11.24) cuando $n = 2$.

Para una región simplemente conexa, la condición $\partial P/\partial y = \partial Q/\partial x$ implica que la integral de línea $\int P\, dx + Q\, dy$ es independiente del camino (teorema 11.11). Como ya hemos observado, si S *no* es simplemente conexo, la condición $\partial P/\partial y = \partial Q/\partial x$ no implica necesariamente la independencia del camino. No obstante, para este caso existe una condición de independencia que se deduce del teorema 11.12.

TEOREMA 11.13. INVARIANCIA DE UNA INTEGRAL DE LÍNEA AL DEFORMAR EL CAMINO. *Sean P y Q derivables con continuidad en un conjunto abierto conexo S del plano, y supongamos que $\partial P/\partial y = \partial Q/\partial x$ en todo S. Sean C_1 y C_2 dos curvas de Jordan regulares a trozos situadas en S y que satisfagan las siguientes condiciones:*

a) *C_2 está en el interior de C_1.*

b) *Los puntos interiores a C_1 que son exteriores a C_2 pertenecen a S.* (*En la figura* 11.19 *se ve un ejemplo.*) *Tenemos entonces*

$$\oint_{C_1} P\,dx + Q\,dy = \oint_{C_2} P\,dx + Q\,dy, \tag{11.25}$$

recorriéndose ambas curvas en el mismo sentido.

Demostración. En las condiciones establecidas, la ecuación (11.24) es aplicable cuando $n = 2$. La región R está constituida por los puntos situados entre las dos curvas C_1 y C_2 y las propias curvas. Puesto que $\partial P/\partial y = \partial Q/\partial x$ en S, el primer miembro de (11.24) es cero y obtenemos (11.25).

Algunas veces el teorema 11.13 se expresa diciendo que si $\partial P/\partial y = \partial Q/\partial x$

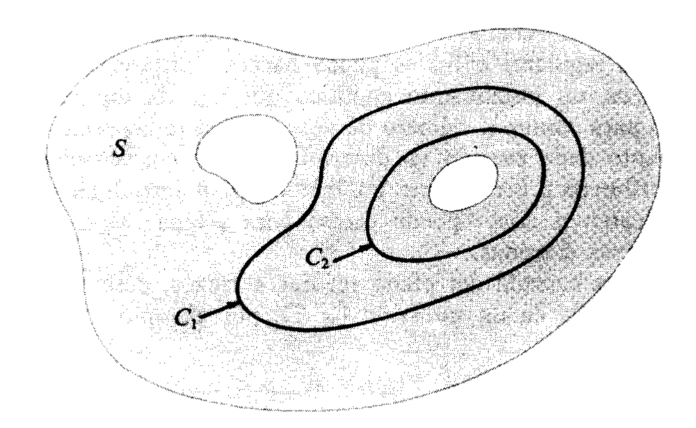


FIGURA 11.19 *Invariancia de una integral de línea al deformar el camino.*

en S el valor de una integral de línea a lo largo de una curva cerrada simple en S no varía si el camino se cambia por deformación en otra curva cerrada simple cualquiera de S, con tal que todas las curvas intermedias que se van obteniendo en la deformación permanezcan dentro de S. Se supone que el conjunto S es conexo y abierto — no es preciso que sea simplemente conexo.

* 11.24 El número de giros

Hemos visto que el valor de una integral de línea depende frecuentemente a la vez de la curva a lo largo de la cual se integra y del sentido en el que dicha curva se recorre. Por ejemplo, la identidad del teorema de Green exige que la integral se tome en el sentido contrario al de las agujas del reloj. En un estudio ri-

guroso del teorema de Green sería necesario describir analíticamente lo que significa recorrer una curva cerrada en el sentido «contrario al de las agujas del reloj». Para ciertas curvas puede hacerse esto estableciendo convenios particulares relativos a la función vectorial $\boldsymbol{\alpha}$ que describe la curva. Por ejemplo, la función vectorial $\boldsymbol{\alpha}$ definida en el intervalo $[0, 2\pi]$ por la ecuación

(11.26) $$\boldsymbol{\alpha}(t) = (a\cos t + x_0)\boldsymbol{i} + (a\operatorname{sen} t + y_0)\boldsymbol{j}$$

describe una circunferencia de radio a y centro en (x_0, y_0). Se dice que esta función describe la circunferencia en sentido *positivo* o *contrario al de las agujas del reloj*. Por otra parte, si reemplazamos t por $-t$ en el segundo miembro de (11.26) obtenemos una nueva función de la que se dice que describe la circunferencia en sentido *negativo* o *de las agujas del reloj*. De este modo hemos dado una descripción completamente analítica del sentido para una circunferencia. Sin embargo, no es tan sencillo hacer lo mismo con una curva cerrada *cualquiera*. Para curvas regulares a trozos puede hacerse introduciendo el concepto del *número de giros*, un instrumento analítico que nos da un método matemáticamente preciso para contar el número de veces que el radio vector $\boldsymbol{\alpha}$ «gira alrededor» de un punto dado cuando va describiendo una curva cerrada dada. En esta sección describiremos brevemente un método para introducir el número de giros. Luego indicaremos cómo puede usarse para asignar sentidos positivos y negativos a las curvas cerradas.

Sea C una curva cerrada del plano regular a trozos, descrita por una función vectorial $\boldsymbol{\alpha}$ definida en un intervalo $[a, b]$, por ejemplo

$$\boldsymbol{\alpha}(t) = X(t)\boldsymbol{i} + Y(t)\boldsymbol{j} \quad \text{si} \quad a \le t \le b.$$

Sea $P_0 = (x_0, y_0)$ un punto no situado en la curva C. El número de giros de $\boldsymbol{\alpha}$ con respecto al punto P_0 se designa por $W(\boldsymbol{\alpha}; P_0)$; se define como el valor de la siguiente integral

(11.27) $$W(\boldsymbol{\alpha}; P_0) = \frac{1}{2\pi}\int_a^b \frac{[X(t) - x_0]Y'(t) - [Y(t) - y_0]X'(t)}{[X(t) - x_0]^2 + [Y(t) - y_0]^2}\, dt.$$

Esto es lo mismo que la integral de línea

(11.28) $$\frac{1}{2\pi}\oint_C \frac{-(y - y_0)\, dx + (x - x_0)\, dy}{(x - x_0)^2 + (y - y_0)^2}.$$

Puede demostrarse que el valor de esta integral siempre es un *entero*, positivo, negativo o nulo. Además, si C es una curva de *Jordan* (curva cerrada *simple*) ese

entero es 0 si P_0 es *exterior* a C y vale $+1$ ó -1 si P_0 es *interior* a C. (Ver figura 11.20.) Asimismo, $W(\boldsymbol{\alpha}; P_0)$ es $+1$ ó -1 para *todo* punto P_0 interior a C.



FIGURA 11.20 *Posibles valores del número de giros de una curva de Jordan C con respecto al punto P_0.*

Esto nos permite definir las orientaciones positiva y negativa para C del siguiente modo: Si el número de giros $W(\boldsymbol{\alpha}; P_0)$ es $+1$ para todo punto P_0 interior a C decimos que $\boldsymbol{\alpha}$ describe C en *sentido positivo* o *contrario al de las agujas del reloj*. Si el número de giros es -1 decimos que $\boldsymbol{\alpha}$ describe C en *sentido negativo* o *de las agujas del reloj*. [En el ejemplo 2 de la sección 10.16, ya se vio un ejemplo de la integral (11.28).]

Para probar que la integral que da el número de giros es siempre $+1$ ó -1 para una curva cerrada simple alrededor de (x_0, y_0), utilizamos el teorema 11.13. Designemos con S la región conexa abierta constituida por todos los puntos del plano excepto (x_0, y_0). La integral de línea (11.28) puede entonces escribirse en la forma $\int_C P\,dx + Q\,dy$, y se comprueba fácilmente que $\partial P/\partial y = \partial Q/\partial x$ en toda la región S. Por consiguiente, si (x_0, y_0) es interior a C, el teorema 11.13 nos dice que podemos reemplazar la curva C por una circunferencia con centro en (x_0, y_0) sin que cambie el valor de la integral. Comprobamos seguidamente que para una circunferencia la integral que da el número de giros es $+1$ ó -1, según que la circunferencia esté orientada positiva o negativamente. Para una circunferencia orientada positivamente podemos utilizar la representación de la ecuación (11.26). En tal caso tenemos

$$X(t) = a\cos t + x_0, \qquad Y(t) = a \operatorname{sen} t + y_0,$$

y el integrando de (11.27) es idénticamente igual a 1. Por tanto obtenemos

$$W(\boldsymbol{\alpha}; P_0) = \frac{1}{2\pi}\int_0^{2\pi} 1\,dt = 1.$$

Mediante un razonamiento análogo encontramos que la integral es -1 cuando C

está orientada negativamente. Esto demuestra que el número de giros es $+1$ ó -1 para una curva cerrada simple que incluye en su interior el punto (x_0, y_0).

* 11.25 Ejercicios

1. Sean $S = \{(x, y) \mid x^2 + y^2 > 0\}$, y

$$P(x, y) = \frac{y}{x^2 + y^2}, \qquad Q(x, y) = \frac{-x}{x^2 + y^2}$$

si $(x,y) \in S$. Sea C una curva de Jordan regular a trozos contenida en S.
a) Si $(0, 0)$ es interior a C, demostrar que la integral de línea $\int_C P\,dx + Q\,dy$ toma el valor $\pm 2\pi$, y razonar cuando el signo es más.
b) Calcular el valor de la integral $\int_C P\,dx + Q\,dy$ cuando $(0, 0)$ es exterior a C.

2. Si $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$ y $r = \|\boldsymbol{r}\|$, sea

$$\boldsymbol{f}(x, y) = \frac{\partial(\log r)}{\partial y}\boldsymbol{i} - \frac{\partial(\log r)}{\partial x}\boldsymbol{j}$$

para $r>0$. Sea C una curva de Jordan regular a trozos situada en el anillo $1<x^2+y^2<25$. Hallar todos los valores posibles de la integral de línea de $\boldsymbol{f}$ a lo largo de C.

3. Una región plana conexa con un solo «agujero» se llama *doblemente conexa.* (El anillo $1 < x^2 + y^2 < 2$ es un ejemplo.) Si P y Q son derivables con continuidad en una región abierta R doblemente conexa, y si $\partial P/\partial y = \partial Q/\partial x$ en todo R. ¿Cuántos valores distintos posibles existen para las integrales $\int_C P\,dx + Q\,dy$ tomadas sobre curvas de Jordan regulares a trozos situadas en R?

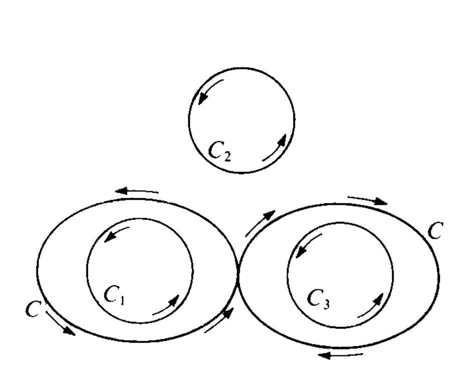


FIGURA 11.21 *Ejercicio 5*

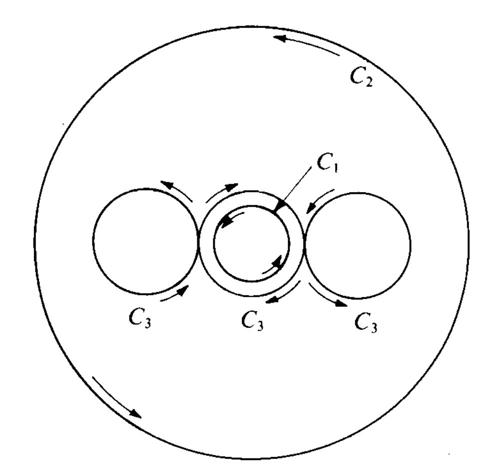


FIGURA 11.22 *Ejercicio 6*

4. Resolver el ejercicio 3 para regiones triplemente conexas, esto es, para regiones planas con dos agujeros solamente.
5. Sean P y Q dos campos escalares con derivadas continuas que satisfacen $\partial P/\partial y = \partial Q/\partial x$ en todo el plano excepto en tres puntos. Sean C_1, C_2, C_3 tres círculos, como muestra la figura 11.21, y sea $I_k = \oint_{C_k} P\,dx + Q\,dy$. Supongamos que $I_1 = 12$, $I_2 = 10$, $I_3 = 15$.

   a) Hallar el valor de $\int_C P\,dx + Q\,dy$, siendo C la curva en forma de ocho.
   b) Dibujar otra curva cerrada Γ a lo largo de la cual $\int P\,dx + Q\,dy = 1$. Indicar en el dibujo el sentido en el que se recorre el camino.
   c) Si $I_1 = 12$, $I_2 = 9$, e $I_3 = 15$, demostrar que no existe camino cerrado alguno Γ a lo largo del cual $\int P\,dx + Q\,dy = 1$.
6. Sea $I_k = \oint_{C_k} P\,dx + Q\,dy$, donde

$$P(x, y) = -y\left[\frac{1}{(x-1)^2 + y^2} + \frac{1}{x^2 + y^2} + \frac{1}{(x+1)^2 + y^2}\right]$$

y

$$Q(x, y) = \frac{x-1}{(x-1)^2 + y^2} + \frac{x}{x^2 + y^2} + \frac{x+1}{(x+1)^2 + y^2}.$$

En la figura 11.22, C_1 es la circunferencia menor, $x^2 + y^2 = \frac{1}{8}$ (con el sentido contrario al de las agujas del reloj), C_2 es la circunferencia mayor, $x^2 + y^2 = 4$ (sentido contrario al de las agujas del reloj), y C_3 es la curva reunión de las tres circunferencias $(x-1)^2 + y^2 = \frac{1}{4}$, $x^2 + y^2 = \frac{1}{4}$, y $(x+1)^2 + y^2 = \frac{1}{4}$ dibujadas en los sentidos que se indican en la figura. Si $I_2 = 6\pi$ e $I_3 = 2\pi$, hallar el valor de I_1.

11.26 Cambio de variables en una integral doble

En la teoría de integración unidimensional el método de sustitución nos permite calcular integrales complicadas transformándolas en otras más sencillas o en tipos que pueden calcularse más fácilmente. El método se basa en la fórmula

$$\int_a^b f(x)\,dx = \int_c^d f[g(t)]g'(t)\,dt, \tag{11.29}$$

donde $a = g(c)$ y $b = g(d)$. En el volumen I se demostró esta fórmula suponiendo que g tiene derivadas continuas en $[c, d]$ y que f es continua en el conjunto de valores que toma $g(t)$ al variar t en el intervalo $[c, d]$.

En dos dimensiones existe análogo problema para las integrales dobles. Se transforma una integral doble de la forma $\iint_S f(x, y)\,dx\,dy$, extendida a una región S en el plano xy, en otra integral doble $\iint_T F(u, v)\,du\,dv$, extendida a una nueva región T del plano uv. Vamos a estudiar a continuación la relación entre las regiones S y T y los integrandos $f(x, y)$ y $F(u, v)$. El método de sustitución en integrales dobles es más laborioso que en las simples debido a que existen dos sustituciones formales a efectuar, una respecto a x y otra respecto a y. Esto signi-

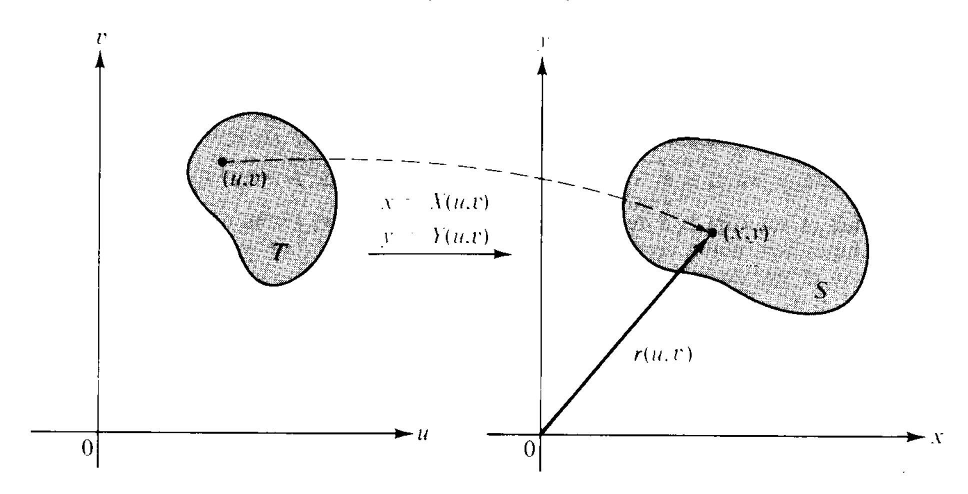


FIGURA 11.23 *Aplicación definida por la ecuación vectorial* $\boldsymbol{r}(u, v) = X(u, v)\boldsymbol{i} + Y(u, v)\boldsymbol{j}$.

fica que en lugar de una función g que aparece en (11.29), tenemos ahora dos funciones, X e Y, que relacionan x, y con u, v del modo siguiente:

$$(11.30) \qquad x = X(u, v), \qquad y = Y(u, v).$$

Geométricamente, puede considerarse que las dos ecuaciones (11.30) definen una «aplicación» que hace corresponder a un punto (u, v) del plano uv, el punto imagen (x, y) del plano xy. Un conjunto T de puntos del plano uv es aplicado en otro conjunto S del xy, como se representa en la figura 11.23. La aplicación también puede expresarse mediante una función vectorial. En el plano xy trazamos el radio vector $\boldsymbol{r}$ que une el origen con un punto genérico (x, y) de S (véase figura 11.23). El vector $\boldsymbol{r}$ depende de u y v y puede considerarse como una función vectorial de dos variables definida por la ecuación

$$(11.31) \qquad \boldsymbol{r}(u, v) = X(u, v)\boldsymbol{i} + Y(u, v)\boldsymbol{j} \quad \text{si} \quad (u, v) \in T.$$

Esta ecuación se llama la *ecuación vectorial* de la aplicación. Como (u, v) recorre puntos de T, el vértice de $\boldsymbol{r}(u, v)$ describe puntos de S.

Algunas veces las dos ecuaciones (11.30) pueden resolverse despejando u y v en función de x e y. Cuando ello es posible podemos expresar el resultado en la forma

$$u = U(x, y), \qquad v = V(x, y).$$

Estas ecuaciones definen una aplicación del plano xy en el plano uv, llamada *aplicación inversa* de la definida por (11.30), ya que transforma los puntos de S

en los de T. Las llamadas *aplicaciones uno a uno* son de especial importancia. Esas transforman puntos *distintos* de T en puntos *distintos* de S; dicho de otro modo, dos puntos distintos de T no se aplican en el mismo punto de S mediante una aplicación uno a uno. Tales aplicaciones establecen una correspondencia uno a uno entre los puntos de T y los correspondientes de S y nos permite (por lo menos teóricamente) regresar de S a T por la aplicación inversa (que, naturalmente, es también uno a uno).

Consideraremos aplicaciones para las que las funciones X e Y son continuas y tienen derivadas parciales continuas $\partial X/\partial u, \partial X/\partial v, \partial Y/\partial u$ y $\partial Y/\partial v$ en S. Para las funciones U y V se hacen análogas hipótesis. Esas restricciones no son en realidad muy rigurosas pues las cumplen la mayoría de las funciones que surgen en la práctica.

La fórmula para la transformación de integrales dobles puede escribirse así:

(11.32) $$\iint_S f(x, y)\, dx\, dy = \iint_T f[X(u, v), Y(u, v)]\, |J(u, v)|\, du\, dv.$$

El factor $J(u, v)$ que aparece en el integrando del segundo miembro desempeña el papel del factor $g'(t)$ de la fórmula (11.29). Este factor se llama *jacobiano* de la aplicación definida por (11.30); es igual al determinante

$$J(u, v) = \begin{vmatrix} \dfrac{\partial X}{\partial u} & \dfrac{\partial Y}{\partial u} \\ \dfrac{\partial X}{\partial v} & \dfrac{\partial Y}{\partial v} \end{vmatrix}.$$

Algunas veces se usa el símbolo $\partial(X, Y)/\partial(u, v)$ en lugar de $J(u, v)$ para representar el jacobiano.

No discutiremos las condiciones más generales en las que es válida la fórmula (11.32). Puede demostrarse (*) que (11.32) es válida si, además de la continuidad de X, Y, U y V antes mencionada, suponemos que la aplicación de T en S es uno a uno y que el jacobiano $J(u, v)$ es distinto de cero. La fórmula también es válida aunque la aplicación no sea uno a uno en un subconjunto de medida nula o el Jacobiano se anule en un conjunto de esta clase.

En la sección 11.30 demostraremos como la fórmula de transformación (11.32) puede deducirse de uno de sus casos particulares, a saber, el caso en que S es un rectángulo y la función f tiene el valor constante 1 en cada punto de S. En este caso particular se transforma en

(11.33) $$\iint_S dx\, dy = \iint_T |J(u, v)|\, du\, dv.$$

(*) Véase el teorema 10-30 del *Análisis matemático* del autor.

Aún en este caso la demostración no es sencilla. En la sección (11.29) se dará una demostración de (11.33) con el auxilio del teorema de Green. En lo que resta de esta sección presentaremos un razonamiento puramente geométrico que justifica la validez de (11.33).

Interpretación geométrica de la ecuación (11.33): Tomemos una región T en el plano uv, como se representa en la figura 11.23, y sea S el conjunto de puntos del plano xy en los que T es aplicado por la función vectorial $\boldsymbol{r}$ dada por (11.31). Introducimos ahora dos nuevas funciones vectoriales $\boldsymbol{V}_1$ y $\boldsymbol{V}_2$ que se obtienen tomando las derivadas parciales de los componentes de $\boldsymbol{r}$ con respecto a u y v, respectivamente. Esto es, definimos

$$\boldsymbol{V}_1 = \frac{\partial \boldsymbol{r}}{\partial u} = \frac{\partial X}{\partial u}\boldsymbol{i} + \frac{\partial Y}{\partial u}\boldsymbol{j} \quad \text{y} \quad \boldsymbol{V}_2 = \frac{\partial \boldsymbol{r}}{\partial v} = \frac{\partial X}{\partial v}\boldsymbol{i} + \frac{\partial Y}{\partial v}\boldsymbol{j}.$$

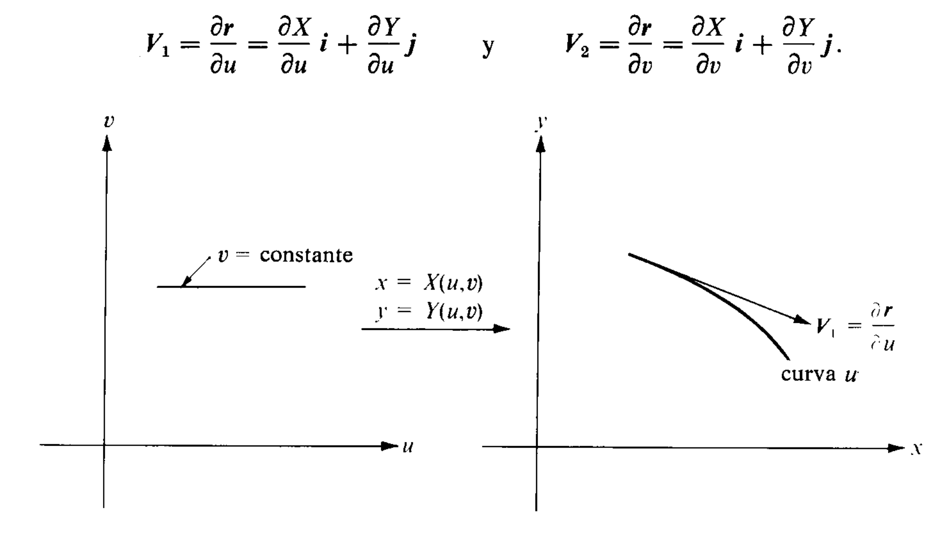


FIGURA 11.24 *Una curva u y el correspondiente vector velocidad.*

Estos vectores pueden interpretarse geométricamente como sigue: Consideremos un segmento rectilíneo horizontal en el plano uv (sobre dicho segmento v es constante). La función vectorial $\boldsymbol{r}$ aplica este segmento sobre una curva (llamada curva u) en el plano xy, como se indica en la figura 11.24. Si consideramos u como un parámetro representando «tiempo», el vector $\boldsymbol{V}_1$ representa la velocidad de la posición $\boldsymbol{r}$ y es por lo tanto tangente a la curva descrita por el extremo de $\boldsymbol{r}$. Del mismo modo, cada vector $\boldsymbol{V}_2$ representa el vector velocidad de una curva v obtenida haciendo u = constante. Una curva u y una curva v pasan por cada punto de la región S.

Consideremos ahora un pequeño rectángulo de dimensiones Δu y Δv, como se presenta en la figura 11.25. Si Δu es la longitud de un pequeño intervalo de tiempo, un punto de una curva u se desplaza, en el tiempo Δu, a lo largo de la

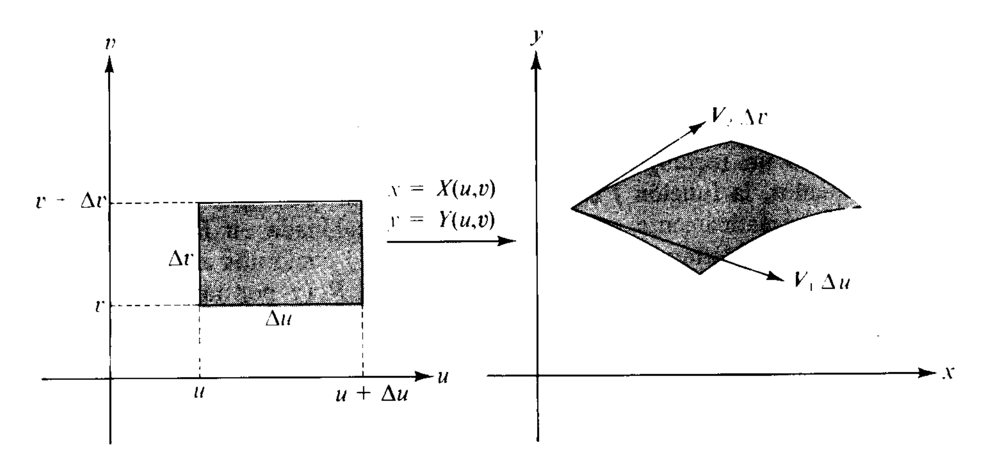


FIGURA 11.25 *La imagen de una región rectangular del plano uv es un paralelogramo curvilíneo en el plano xy.*

misma una distancia aproximadamente igual al producto $\|V_1\| \Delta u$ (ya que $\|V_1\|$ representa la velocidad y Δu el tiempo). Análogamente, en el tiempo Δv un punto de una curva v se desplaza a lo largo de ella una distancia aproximadamente igual a $\|V_2\| \Delta v$. Por consiguiente, la región rectangular de dimensiones Δu y Δv del plano uv se transforma en una porción del plano xy que es casi un paralelogramo, cuyos lados son los vectores $V_1 \Delta u$ y $V_2 \Delta v$ como se ve en la figura 11.25. El área de ese paralelogramo es el valor del producto exterior de los dos vectores $V_1 \Delta u$ y $V_2 \Delta v$; éste es igual a

$$\|(V_1 \Delta u) \times (V_2 \Delta v)\| = \|V_1 \times V_2\| \Delta u \Delta v .$$

Si calculamos el producto $V_1 \times V_2$ en función de los componentes de V_1 y V_2 encontramos

$$V_1 \times V_2 = \begin{vmatrix} i & j & k \\ \dfrac{\partial X}{\partial u} & \dfrac{\partial Y}{\partial u} & 0 \\ \dfrac{\partial X}{\partial v} & \dfrac{\partial Y}{\partial v} & 0 \end{vmatrix} = \begin{vmatrix} \dfrac{\partial X}{\partial u} & \dfrac{\partial Y}{\partial u} \\ \dfrac{\partial X}{\partial v} & \dfrac{\partial Y}{\partial v} \end{vmatrix} k = J(u, v)k .$$

Luego la magnitud de $V_1 \times V_2$ es exactamente $J(u, v)$ y el área del paralelogramo curvilíneo de la figura 11.25 es casi igual a $J(u, v) \Delta u \Delta v$.

Si $J(u, v) = 1$ para todos los puntos de T, el «paralelogramo» tiene la misma área que el rectángulo y la aplicación conserva las áreas. Si no es así, para obtener

el área del paralelogramo debemos multiplicar el área del rectángulo por $|J(u, v)|$. Esto sugiere que el jacobiano puede imaginarse como un «factor de ampliación» de áreas.

Sea ahora P una partición de un rectángulo R que contenga toda la región T y consideremos un rectángulo parcial de P de dimensiones Δu y Δv. Si Δu y Δv son pequeños, la función J es casi constante en el rectángulo parcial y por lo tanto J actúa prácticamente como una función escalonada en R. (Asignamos a J el valor cero en el exterior de T.) Si consideramos J como una auténtica función escalonada, entonces la integral doble de $|J|$ sobre R (y por tanto sobre T) es una suma de productos de la forma $|J(u, v)|\, \Delta u\, \Delta v$ y las anteriores consideraciones sugieren que esta suma es aproximadamente igual al área de S, que conocemos como la integral doble $\iint_S dx\, dy$.

Esta discusión geométrica, que tan sólo sugiere por qué puede esperarse la validez de una fórmula como la (11.33), puede constituir la base de una demostración rigurosa, pero los detalles son largos e intrincados. Como antes se ha dicho, se dará en un capítulo posterior una demostración de (11.33) con fundamento totalmente distinto.

Si $J(u, v) = 0$ en un punto (u, v), los dos vectores V_1 y V_2 son paralelos (ya que su producto exterior es el vector nulo) y el paralelogramo degenera. Tales puntos se llaman *puntos singulares* de la aplicación. Como ya hemos mencionado, la fórmula de transformación (11.32) también es válida con tal que exista un número finito de puntos singulares o, más general, cuando tales puntos formen un conjunto de medida nula. Esto es lo que ocurre en todas las aplicaciones que usaremos. En la sección siguiente empleamos la fórmula (11.32) en dos ejemplos importantes.

11.27 Casos particulares de la fórmula de transformación

EJEMPLO 1. *Coordenadas polares.* En este caso escribimos r y θ en lugar de u y v y definimos la aplicación con las ecuaciones

$$x = r\cos\theta, \qquad y = r\,\mathrm{sen}\,\theta.$$

Esto es, $X(r, \theta) = r\cos\theta$ e $Y(r, \theta) = r\,\mathrm{sen}\,\theta$. Para conseguir una aplicación uno a uno mantenemos $r > 0$ y θ en un intervalo de la forma $\theta_0 \leq \theta < \theta_0 + 2\pi$. Por ejemplo, la aplicación es uno a uno en cualquier subconjunto del rectángulo $(0, a] \times [0, 2\pi)$ del plano $r\theta$. El jacobiano de esta aplicación es:

$$J(r, \theta) = \begin{vmatrix} \dfrac{\partial X}{\partial r} & \dfrac{\partial Y}{\partial r} \\[2ex] \dfrac{\partial X}{\partial \theta} & \dfrac{\partial Y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & \mathrm{sen}\,\theta \\ -r\,\mathrm{sen}\,\theta & r\cos\theta \end{vmatrix} = r(\cos^2\theta + \mathrm{sen}^2\theta) = r\,.$$

Luego la fórmula de transformación (11.32) toma la forma

$$\iint_S f(x, y)\, dx\, dy = \iint_T f(r \cos \theta, r \operatorname{sen} \theta)\, r\, dr\, d\theta .$$

Las curvas r son rectas por el origen y las curvas θ son círculos concéntricos con el origen. La imagen de un rectángulo en el plano $r\theta$ es un «paralelogramo» en el xy limitado por dos radios y dos arcos de círculo, como se ve en la figura 11.26.

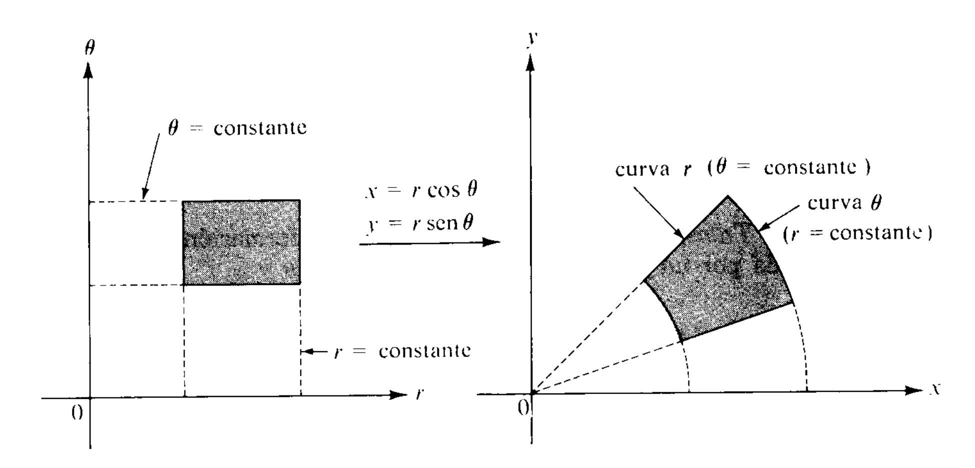


FIGURA 11.26 *Transformación por coordenadas polares.*

El jacobiano se anula cuando $r = 0$, pero esto no afecta la validez de la fórmula de transformación porque el conjunto de puntos con $r = 0$ tiene contenido o medida nula.

Puesto que $V_1 = \cos\theta\, \boldsymbol{i} + \operatorname{sen}\theta\, \boldsymbol{j}$, tenemos $\|V_1\| = 1$, así que no existe distorsión de distancias a lo largo de las curvas r. Por otra parte,

$$V_2 = -r \operatorname{sen}\theta\, \boldsymbol{i} + r\cos\theta\, \boldsymbol{j}, \qquad \|V_2\| = r,$$

de modo que las distancias a lo largo de las curvas θ quedan multiplicadas por el factor r.

Las coordenadas polares son particularmente convenientes cuando la región de integración tiene fronteras a lo largo de las cuales r o θ es constante. Por ejemplo, consideremos la integral para calcular el volumen de un octante de esfera de radio a,

$$\iint_S \sqrt{a^2 - x^2 - y^2}\, dx\, dy ,$$

en la que la región S es el primer cuadrante del disco circular $x^2 + y^2 \leqslant a^2$. En polares la integral se transforma así

$$\iint_T \sqrt{a^2 - r^2}\, r\, dr\, d\theta\,,$$

donde la región de integración T es ahora un rectángulo $[0, a] \times [0, \frac{1}{2}\pi]$. Integrando primero respecto a θ y luego respecto a r obtenemos:

$$\iint_T \sqrt{a^2 - r^2}\, r\, dr\, d\theta = \frac{\pi}{2}\int_0^a r\sqrt{a^2 - r^2}\, dr = \frac{\pi}{2} \cdot \frac{(a^2 - r^2)^{3/2}}{-3}\bigg|_0^a = \frac{\pi a^3}{6}\,.$$

Éste es el mismo resultado que se obtuvo antes, pero en polares resulta el cálculo más cómodo y breve.

EJEMPLO 2. *Transformaciones lineales.* Una transformación lineal es una aplicación definida por un par de ecuaciones de la forma

(11.34) $$x = Au + Bv, \qquad y = Cu + Dv,$$

donde A, B, C, D son constantes dadas. El jacobiano es:

$$J(u, v) = AD - BC\,,$$

y para asegurar la existencia de la transformación inversa suponemos $AD - BC \neq 0$. Esto nos garantiza que podamos resolver las ecuaciones lineales (11.34) respecto a u y v en función de x e y.

Las transformaciones lineales transforman rectas en rectas. Por lo tanto, la imagen de un rectángulo en el plano uv es un cuadrilátero en el xy, y su área es la del rectángulo multiplicada por el factor $|J(u, v)| = |AD - BC|$. La fórmula de transformación (11.32) se convierte en esta otra

$$\iint_S f(x, y)\, dx\, dy = |AD - BC| \iint_T f(Au + Bv, Cu + Dv)\, du\, dv\,.$$

A continuación damos un ejemplo en el que se pone de manifiesto la utilidad de un cambio lineal de variable. Consideremos la integral

$$\iint_S e^{(y-x)/(y+x)}\, dx\, dy\,,$$

en la que S es el triángulo determinado por la recta $x + y = 2$ y los dos ejes

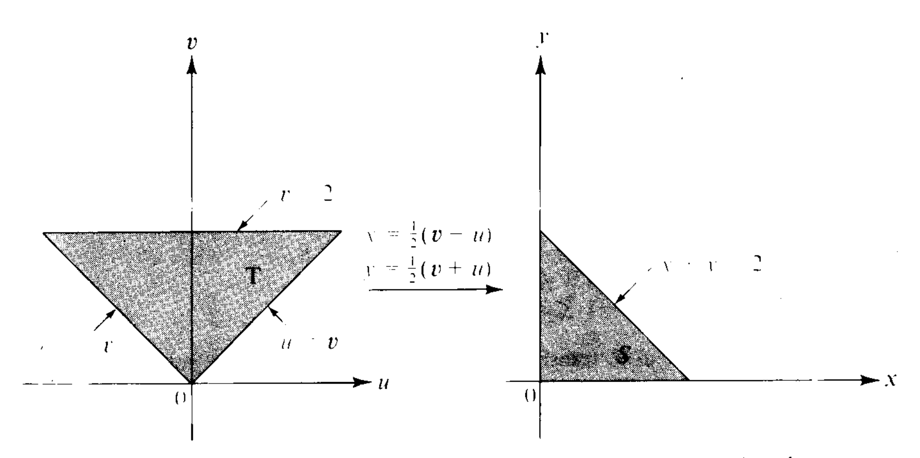


FIGURA 11.27 *Aplicación definida por una transformación lineal*

coordenados. (Véase figura 11.27.) La presencia de $y - x$ e $y + x$ en el integrando sugiere el cambio de variable

$$u = y - x, \qquad v = y + x.$$

Resolviendo respecto a x e y encontramos

$$x = \frac{v - u}{2} \qquad \text{y} \qquad y = \frac{v + u}{2}.$$

El jacobiano es $J(u, v) = -\frac{1}{2}$. Para encontrar la imagen T de S en el plano uv observemos que las rectas $x = 0$ e $y = 0$ se transforman en las rectas $u = v$ y $u = -v$, respectivamente; la recta $x + y = 2$ se transforma en la recta $v = 2$. Los puntos interiores a S satisfacen $0 < x + y < 2$ y se transforman en puntos de T que satisfacen $0 < v < 2$. Por lo tanto, la nueva región de integración T es una región triangular, como se ve en la figura 11.27. La integral doble que consideramos se transforma así

$$\iint_S e^{(y-x)/(y+x)}\, dx\, dy = \frac{1}{2} \iint_T e^{u/v}\, du\, dv.$$

Integrando primero respecto a u encontramos

$$\frac{1}{2} \iint_T e^{u/v}\, du\, dv = \frac{1}{2} \int_0^2 \left[\int_{-v}^{v} e^{u/v}\, du \right] dv = \frac{1}{2} \int_0^2 v\left(e - \frac{1}{e}\right) dv = e - \frac{1}{e}.$$

11.28 Ejercicios

En cada uno de los ejercicios del 1 al 5, dibujar la región S y expresar la integral doble $\iint_S f(x, y)\, dx\, dy$ como una integral reiterada en coordenadas polares.

1. $S = \{(x, y) \mid x^2 + y^2 \le a^2\}$, donde $a > 0$.
2. $S = \{(x, y) \mid x^2 + y^2 \le 2x\}$.
3. $S = \{(x, y) \mid a^2 \le x^2 + y^2 \le b^2\}$, donde $0 < a < b$.
4. $S = \{(x, y) \mid 0 \le y \le 1 - x, 0 \le x \le 1\}$.
5. $S = \{(x, y) \mid x^2 \le y \le 1, -1 \le x \le 1\}$.

En los ejercicios del 6 al 9, transformar la integral a coordenadas polares y calcular su valor. (La letra a representa una constante positiva.)

6. $\int_0^{2a}\left[\int_0^{\sqrt{2ax-x^2}}(x^2 + y^2)\, dy\right] dx$.
7. $\int_0^{a}\left[\int_0^{x}\sqrt{x^2 + y^2}\, dy\right] dx$.
8. $\int_0^{1}\left[\int_{x^2}^{x}(x^2 + y^2)^{-1/2}\, dy\right] dx$.
9. $\int_0^{a}\left[\int_0^{\sqrt{a^2-y^2}}(x^2 + y^2)\, dx\right] dy$.

En los ejercicios del 10 al 13, transformar cada una de las integrales dadas en una o más integrales reiteradas en polares.

10. $\int_0^{1}\left[\int_0^{1} f(x, y)\, dy\right] dx$.
11. $\int_0^{2}\left[\int_x^{x\sqrt{3}} f(\sqrt{x^2 + y^2})\, dy\right] dx$.
12. $\int_0^{1}\left[\int_{1-x}^{\sqrt{1-x^2}} f(x, y)\, dy\right] dx$.
13. $\int_0^{1}\left[\int_0^{x^2} f(x, y)\, dy\right] dx$.

14. Utilizar una conveniente transformación lineal para calcular la integral doble

$$\iint_S (x - y)^2 \operatorname{sen}^2 (x + y)\, dx\, dy$$

donde S es el paralelogramo con vértices $(\pi, 0)$, $(2\pi, \pi)$, $(\pi, 2\pi)$, $(0, \pi)$.

15. Los vértices de un paralelogramo S del plano xy son los puntos $(0, 0)$, $(2, 10)$, $(3, 17)$ y $(1, 7)$.
a) Hallar una transformación lineal $u = ax + by$, $v = cx + dy$, que aplique S en un rectángulo R del plano uv con vértices opuestos en $(0, 0)$ y $(4, 2)$. El vértice $(2, 10)$ deberá aplicarse en un punto del eje u.
b) Calcular la integral doble $\iint_S xy\, dx\, dy$ transformándola en una integral equivalente en el rectángulo R del apartado a).

16 Si $r > 0$, pongamos $I(r) = \int_{-r}^{r} e^{-u^2}\, du$.
a) Demostrar que $I^2(r) = \iint_R e^{-(x^2+y^2)} dx\, dy$, en donde R es el cuadrado $R = [-r, r] \times [-r, r]$.
b) Si C_1 y C_2 son los discos circulares inscrito y circunscrito a R, demostrar que

$$\iint_{C_1} e^{-(x^2+y^2)}\, dx\, dy < I^2(r) < \iint_{C_2} e^{-(x^2+y^2)}\, dx\, dy .$$

c) Expresar las integrales a lo largo de C_1 y C_2 en coordenadas polares y aplicar b) para deducir que $I(r) \to \sqrt{\pi}$ cuando r $\to \infty$. Esto demuestra que $\int_0^\infty e^{-u^2}\,du = \sqrt{\pi}/2$.
d) Aplicar el apartado c) para deducir que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, siendo Γ la función gamma.

17. Considerar la aplicación definida por las ecuaciones

$$x = u + v, \qquad y = v - u^2.$$

a) Calcular el jacobiano $J(u, v)$.
b) Un triángulo T en el plano uv tiene vértices $(0,0)$, $(2,0)$, $(0,2)$. Representar, mediante un dibujo, la imagen S en el plano xy.
c) Calcular el área de S mediante una integral doble extendida a S y también mediante otra integral doble extendida a T.
d) Calcular $\iint_S (x - y + 1)^{-2}\,dx\,dy$.

18. Considerar la aplicación definida por las dos ecuaciones $x = u^2 - v^2$, $y = 2uv$.

a) Calcular el jacobiano $J(u, v)$.
b) Sea T el rectángulo en el plano uv con vértices $(1,1)$, $(2,1)$, $(2,3)$, $(1,3)$. Representar, mediante un dibujo, la imagen S en el plano xy.
c) Calcular la integral doble $\iint_C xy\,dx\,dy$ haciendo el cambio de variable $x = u^2 - v^2$, $y = 2uv$, donde $C = \{(x, y) \mid x^2 + y^2 \leq 1\}$.

19. Calcular la integral doble

$$I(p, r) = \iint_R \frac{dx\,dy}{(p^2 + x^2 + y^2)^p}$$

sobre el disco circular $R = \{(x, y) \mid x^2 + y^2 \leq r^2\}$. Determinar los valores de p para los que $I(p, r)$ tiene límite cuando $r \to +\infty$.

En los ejercicios 20 al 22, establecer las igualdades que se dan mediante la introducción de un conveniente cambio de variable en cada caso.

20. $\iint_S f(x + y)\,dx\,dy = \int_{-1}^{1} f(u)\,du$, donde $S = \{(x, y) \mid |x| + |y| \leq 1\}$.

21. $\iint_S f(ax + by + c)\,dx\,dy = 2\int_{-1}^{1} \sqrt{1 - u^2}\, f(u\sqrt{a^2 + b^2} + c)\,du$,

donde $S = \{(x, y) \mid x^2 + y^2 \leq 1\}$ y $a^2 + b^2 \neq 0$.

22. $\iint_S f(xy)\,dx\,dy = \log 2 \int_1^2 f(u)\,du$, siendo S la región del primer cuadrante limitada por las curvas $xy = 1$, $xy = 2$, $y = x$, $y = 4x$.

11.29 Demostración de la fórmula de transformación en un caso particular

Como ya se ha indicado antes, la fórmula de transformación

$$\iint_S f(x, y)\, dx\, dy = \iint_T f[X(u, v), Y(u, v)]\, |J(u, v)|\, du\, dv \tag{11.35}$$

puede deducirse como una consecuencia del caso particular en el que S es un rectángulo y f es idénticamente igual a 1. En este caso la fórmula se simplifica, reduciéndose a

$$\iint_R dx\, dy = \iint_{R^*} |J(u, v)|\, du\, dv. \tag{11.36}$$

Aquí R representa un rectángulo en el plano xy y R^* representa su imagen en el plano uv (véase figura 11.28) obtenida con una aplicación uno a uno

$$u = U(x, y), \qquad v = V(x, y).$$

La aplicación inversa viene dada por

$$x = X(u, v), \qquad y = Y(u, v),$$

FIGURA 11.28 *Ley de transformación para integrales dobles deducida del teorema de Green.*

y $J(u, v)$ es el determinante jacobiano

$$J(u, v) = \begin{vmatrix} \dfrac{\partial X}{\partial u} & \dfrac{\partial X}{\partial v} \\ \dfrac{\partial Y}{\partial u} & \dfrac{\partial Y}{\partial v} \end{vmatrix}.$$

En esta sección utilizamos el teorema de Green para demostrar (11.36), y en la siguiente deducimos la fórmula más general (11.35) a partir del caso particular (11.36).

Para la demostración suponemos que las funciones X e Y tienen derivadas parciales segundas continuas y que el jacobiano nunca se anula en R^*. Entonces $J(u, v)$ es siempre positivo o siempre negativo. El significado de este signo de $J(u, v)$ es que cuando un punto (x, y) describe la frontera de R en el sentido contrario al de las agujas del reloj, el punto imagen (u, v) describe la frontera de R^* en el mismo sentido si $J(u, v)$ es positivo y en sentido contrario si $J(u, v)$ es negativo. En la demostración supondremos que $J(u, v) > 0$.

La idea de la demostración es la de expresar cada integral doble de (11.36) como una integral de línea usando el teorema de Green. Comprobamos luego la igualdad de las dos integrales de línea expresándolas en forma paramétrica.

Comenzamos con la integral doble en el plano xy, escribiendo

$$\iint_R dx\, dy = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy,$$

en donde $Q(x, y) = x$ y $P(x, y) = 0$. Según el teorema de Green esta integral doble es igual a la integral de línea

$$\int_C P\, dx + Q\, dy = \int_C x\, dy. \tag{11.37}$$

Aquí C es la frontera de R recorrida en el sentido contrario al de las agujas del reloj.

Análogamente, transformamos la integral doble en el plano uv en una integral de línea a lo largo de la frontera C^* de R^*. El integrando, $J(u, v)$, puede escribirse como sigue:

$$\begin{aligned} J(u, v) &= \frac{\partial X}{\partial u}\frac{\partial Y}{\partial v} - \frac{\partial X}{\partial v}\frac{\partial Y}{\partial u} = \frac{\partial X}{\partial u}\frac{\partial Y}{\partial v} + X\frac{\partial^2 Y}{\partial u\, \partial v} - X\frac{\partial^2 Y}{\partial v\, \partial u} - \frac{\partial X}{\partial v}\frac{\partial Y}{\partial u} \\ &= \frac{\partial}{\partial u}\left(X\frac{\partial Y}{\partial v}\right) - \frac{\partial}{\partial v}\left(X\frac{\partial Y}{\partial u}\right). \end{aligned}$$

Aplicando el teorema de Green a la integral doble sobre R^* encontramos

$$\iint_{R^*} J(u, v)\, du\, dv = \int_{C^*} \left(X \frac{\partial Y}{\partial u}\, du + X \frac{\partial Y}{\partial v}\, dv \right).$$

Por consiguiente, para completar la demostración de (11.36) necesitamos tan sólo comprobar que

$$\int_C x\, dy = \int_{C^*} \left(X \frac{\partial Y}{\partial u}\, du + X \frac{\partial Y}{\partial v}\, dv \right).$$

Introducimos una parametrización de C^* y utilizamos ésta para hallar una representación de C. Supongamos que C^* está descrita por una función $\boldsymbol{\alpha}$ definida en un intervalo $[a, b]$, por ejemplo

$$\boldsymbol{\alpha}(t) = U(t)\boldsymbol{i} + V(t)\boldsymbol{j}.$$

Llamemos

$$\boldsymbol{\beta}(t) = X[U(t), V(t)]\boldsymbol{i} + Y[U(t), V(t)]\boldsymbol{j}.$$

Entonces cuando t varía en el intervalo $[a, b]$, el vector $\boldsymbol{\alpha}(t)$ describe la curva C^* y $\boldsymbol{\beta}(t)$ la C. Según la regla de la cadena, la derivada de $\boldsymbol{\beta}$ viene dada por

$$\boldsymbol{\beta}'(t) = \left[\frac{\partial X}{\partial u} U'(t) + \frac{\partial X}{\partial v} V'(t) \right] \boldsymbol{i} + \left[\frac{\partial Y}{\partial u} U'(t) + \frac{\partial Y}{\partial v} V'(t) \right] \boldsymbol{j}.$$

Luego

$$\int_C x\, dy = \int_a^b X[U(t), V(t)] \left(\frac{\partial Y}{\partial u} U'(t) + \frac{\partial Y}{\partial v} V'(t) \right) dt.$$

La última integral sobre $[a, b]$ se obtiene también parametrizando la integral de línea sobre C^* en (11.37). Por consiguiente las dos integrales de línea (11.37) son iguales, lo que demuestra (11.36).

11.30 Demostración de la fórmula de transformación en el caso general

En esta sección deducimos la fórmula general de transformación

$$(11.38) \qquad \iint_S f(x, y)\, dx\, dy = \iint_T f[X(u, v), Y(u, v)]\, |J(u, v)|\, du\, dv$$

a partir del caso particular tratado en la sección anterior,

$$(11.39) \qquad \iint_R dx\,dy = \iint_{R^*} |J(u,v)|\,du\,dv,$$

en donde R es un rectángulo y R^* su imagen en el plano uv.

Demostramos ante todo que

$$(11.40) \qquad \iint_R s(x,y)\,dx\,dy = \iint_{R^*} s[X(u,v),\,Y(u,v)]\,|J(u,v)|\,du\,dv,$$

siendo s una función escalonada cualquiera definida en R. A tal fin, sean P una partición de R en mn subrectángulos R_{ij} de dimensiones Δx_i y Δy_j, y c_{ij} el valor constante que toma s en el subrectángulo abierto R_{ij}. Aplicando (11.39) al rectángulo R_{ij} encontramos

$$\Delta x_i\,\Delta y_j = \iint_{R_{ij}} dx\,dy = \iint_{R_{ij}^*} |J(u,v)|\,du\,dv.$$

Multiplicando ambos miembros por c_{ij} y sumando respecto a los índices i, j obtenemos

$$(11.41) \qquad \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}\,\Delta x_i\,\Delta y_j = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij} \iint_{R_{ij}^*} |J(u,v)|\,du\,dv.$$

Puesto que s es una función escalonada, esto es lo mismo que

$$(11.42) \qquad \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}\,\Delta x_i\,\Delta y_j = \sum_{i=1}^{n}\sum_{j=1}^{m} \iint_{R_{ij}^*} s[X(u,v),\,Y(u,v)]\,|J(u,v)|\,du\,dv.$$

Utilizando la propiedad aditiva de las integrales dobles vemos que (11.42) es lo mismo que (11.40). Así pues, (11.40) es una consecuencia de (11.39).

Seguidamente demostramos que la función escalonada s de (11.40) puede reemplazarse por cualquier función f para la cual existan ambos miembros de (11.40). Sea f integrable en un rectángulo R y elijamos funciones escalonadas s y t que satisfagan las desigualdades

$$(11.43) \qquad s(x,y) \le f(x,y) \le t(x,y),$$

para todos los puntos (x, y) de R. Tenemos entonces

$$(11.44) \qquad s[X(u,v),\,Y(u,v)] \le f[X(u,v),\,Y(u,v)] \le t[X(u,v),\,Y(u,v)]$$

para todo punto (u, v) en la imagen R^*. Por brevedad, escribimos $S(u, v)$ en lugar de $s[X(u, v), Y(u, v)]$ y definimos $F(u, v)$ y $T(u, v)$ de modo parecido. Multiplicando las desigualdades de (11.44) por $|J(u, v)|$ e integrando en R^* obtenemos

$$\iint_{R^*} S(u, v)\,|J(u, v)|\,du\,dv \le \iint_{R^*} F(u, v)\,|J(u, v)|\,du\,dv \le \iint_{R^*} T(u, v)\,|J(u, v)|\,du\,dv\,.$$

Debido a (11.40), las desigualdades anteriores son las mismas que

$$\iint_{R} s(x, y)\,dx\,dy \le \iint_{R^*} F(u, v)\,|J(u, v)|\,du\,dv \le \iint_{R} t(x, y)\,dx\,dy\,.$$

Por consiguiente, $\iint_{R^*} F(u, v)\,|J(u, v)|\,du\,dv$ es un número comprendido entre las integrales $\iint_R s(x, y)\,dx\,dy$ e $\iint_R t(x, y)\,dx\,dy$ para todo par de funciones escalonadas s y t elegidas que satisfagan (11.43). Ya que f es integrable, esto implica que

$$\iint_{R} f(x, y)\,dx\,dy = \iint_{R^*} F(u, v)\,|J(u, v)|\,du\,dv$$

y por tanto (11.38) es válida para funciones integrables definidas sobre rectángulos.

Una vez probada la validez de (11.38) para rectángulos podemos fácilmente extenderla a regiones S más generales por el procedimiento usual de elegir un rectángulo R que contenga a S y considerar en lugar de f otra función $\tilde{f}$ que coincide con f en S y vale 0 fuera de S. Entonces observamos que

$$\iint_{S} f = \iint_{R} \tilde{f} = \iint_{R^*} \tilde{f}[X(u, v), Y(u, v)]\,|J(u, v)|\,du\,dv = \iint_{T} F(u, v)\,|J(u, v)|\,du\,dv$$

lo cual demuestra que (11.38) es, en realidad, una consecuencia de (11.39).

11.31 Extensiones a un número mayor de dimensiones

El concepto de integral múltiple puede extenderse del espacio de dos dimensiones al de n dimensiones para cualquier $n \geq 3$. Puesto que el desarrollo es completamente análogo al caso $n = 2$ tan sólo apuntamos los principales resultados.

El integrando es un campo escalar f definido y acotado en un conjunto S

del n-espacio. La integral de f en S, llamada integral n-múltiple, se representa mediante los símbolos.

$$\int \cdots \int_S f, \qquad \text{o} \qquad \int \cdots \int_S f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n,$$

con n signos de integral, o más sencillamente con un signo integral, $\int_S f(\boldsymbol{x})\, d\boldsymbol{x}$, en donde $\boldsymbol{x} = (x_1, \ldots, x_n)$. Cuando $n = 3$ escribimos (x, y, z) en lugar de (x_1, x_2, x_3) y ponemos para las integrales triples

$$\iiint_S f, \qquad \text{o} \qquad \iiint_S f(x, y, z)\, dx\, dy\, dz\,.$$

Definimos primero la integral n-múltiple para una función escalonada definida en un intervalo *n-dimensional*. Recordemos que un intervalo cerrado n-dimensional $[\boldsymbol{a}, \boldsymbol{b}]$ es el producto cartesiano de n intervalos unidimensionales cerrados $[a_k, b_k]$, en donde $\boldsymbol{a} = (a_1, \ldots, a_n)$ y $\boldsymbol{b} = (b_1, \ldots, b_n)$. Un intervalo n-dimensional abierto $(\boldsymbol{a}, \boldsymbol{b})$ es el producto cartesiano de n intervalos abiertos (a_k, b_k). El volumen de $[\boldsymbol{a}, \boldsymbol{b}]$ o de $(\boldsymbol{a}, \boldsymbol{b})$ se define como el producto de las longitudes de los intervalos componentes

$$(b_1 - a_1) \cdots (b_n - a_n)\,.$$

Si $P_1, \ldots, P_n$ son particiones de $[a_1, b_1], \ldots, [a_n, b_n]$, respectivamente, el producto cartesiano $P = P_1 \times \ldots \times P_n$ será una partición de $[\boldsymbol{a}, \boldsymbol{b}]$. Una función f definida en $[\boldsymbol{a}, \boldsymbol{b}]$ se llama escalonada si es constante en cada uno de los subintervalos abiertos determinados por una cierta partición P. Entonces la integral n-múltiple de una tal función escalonada está definida por la fórmula

$$\int \cdots \int_{[\boldsymbol{a},\boldsymbol{b}]} f = \sum_i c_i v_i\,,$$

en donde c_i es el valor constante que toma f en el i-ésimo subintervalo abierto y v_i es el volumen del i-ésimo subintervalo. La suma es una suma finita extendida a todos los subintervalos de P.

A partir de la integral n-múltiple para funciones escalonadas, definimos la integral para funciones acotadas más generales definidas en intervalos, siguiendo el método corriente. Sean s y t funciones escalonadas tales que $s \leqslant f \leqslant t$ en $[\boldsymbol{a}, \boldsymbol{b}]$. Si existe un número I y sólo uno tal que

$$\int \cdots \int_{[\boldsymbol{a},\boldsymbol{b}]} s \leq I \leq \int \cdots \int_{[\boldsymbol{a},\boldsymbol{b}]} t$$

cualesquiera que sean s y t que satisfagan $s \leqslant f \leqslant t$, entonces se dice que f es integrable en $[\boldsymbol{a}, \boldsymbol{b}]$, y el número I es la integral n-múltiple de f,

$$I = \int \cdots \int_{[a,b]} f.$$

Como en el caso bi-dimensional, la integral existe si f es continua en $[\boldsymbol{a}, \boldsymbol{b}]$. También existe si f es acotada en $[\boldsymbol{a}, \boldsymbol{b}]$ y si el conjunto de discontinuidad de f tiene contenido n-dimensional nulo. Un conjunto acotado S tiene contenido n-dimensional nulo si para todo $\epsilon > 0$ existe una colección finita de intervalos n-dimensionales cuya reunión cubre a S y tal que la suma de sus volúmenes no supera a ϵ.

Para definir la integral n-dimensional de una función acotada f sobre un conjunto acotado S más general, consideramos una nueva función $\tilde{f}$ extensión de f y que coincide con ella en S y vale 0 fuera de S; la integral de f sobre S se define como la integral de $\tilde{f}$ sobre un intervalo que contenga a S.

Algunas integrales múltiples pueden calcularse mediante integrales reiteradas de dimensión inferior. Por ejemplo, supongamos que S es un conjunto en el espacio tri-dimensional definido así:

$$S = \{(x, y, z) \mid (x, y) \in Q \quad \text{y} \quad \varphi_1(x, y) \leq z \leq \varphi_2(x, y)\}, \tag{11.45}$$

siendo Q una región bi-dimensional, llamada proyección de S en el plano x, y, y φ_1 y φ_2 funciones continuas en S. (En la figura 11.29 se muestra un ejemplo.) Los conjuntos de este tipo están limitados por dos superficies de ecuaciones cartesianas $z = \varphi_1(x, y)$ y $z = \varphi_2(x, y)$ y (a veces) una porción de superficie cilíndrica engendrada por una recta que se desplaza paralelamente al eje z siguiendo la frontera de Q. Las rectas paralelas al eje z tienen comunes con ese conjunto segmentos rectilíneos que unen la superficie superior con la inferior. Si f es continua en el interior de S tenemos la fórmula de iteración

$$\iiint_S f(x, y, z)\, dx\, dy\, dz = \iint_Q \left[\int_{\varphi_1(x,y)}^{\varphi_2(x,y)} f(x, y, z)\, dz\right] dx\, dy. \tag{11.46}$$

Esto es, para x e y fijos, la primera integración se efectúa respecto a z desde la superficie frontera inferior hasta la superior. Esto reduce el cálculo a una integral doble sobre la proyección Q, que puede tratarse con los métodos antes expuestos.

Existen otros dos tipos de conjuntos análogos a los descritos por (11.45) en los que los ejes x e y desempeñan el papel del eje z, con las proyecciones en los planos yz o xz, respectivamente. Las integrales triples sobre tales conjuntos pueden calcularse por iteración, con fórmulas análogas a (11.46). La mayoría de los

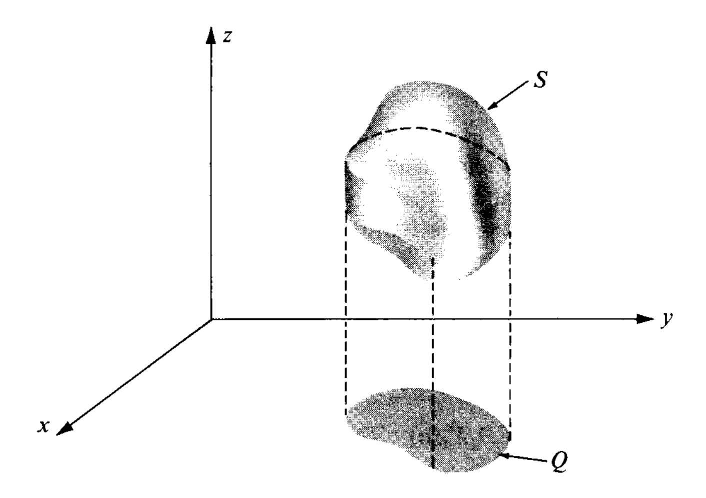


FIGURA 11.29 *Un sólido S y su proyección Q en el plano xy.*

conjuntos tri-dimensionales que encontraremos son de uno de los tres tipos mencionados o pueden descomponerse en un número finito de fragmentos, cada uno de los cuales es de uno de esos tipos.

Existen muchas fórmulas de integración reiterada para las integrales n-múltiples cuando $n > 3$. Por ejemplo, si Q es un intervalo k-dimensional y R un intervalo m-dimensional, entonces una integral $(m + k)$-múltiple sobre $Q \times R$ es la iteración de una integral m-múltiple y una integral k-múltiple.

$$\int_{Q\times R}\!\cdots\!\int f = \int_{Q}\!\cdots\!\int \left[\int_{R}\!\cdots\!\int f\,dx_1\cdots dx_m\right] dx_{m+1}\cdots dx_{m+k},$$

con tal que existan todas las integrales múltiples que aparezcan en el cálculo. Luego veremos en este capítulo como se usan las integrales múltiples iteradas al calcular el volumen de una esfera n-dimensional.

11.32 Cambio de variables en una integral n-múltiple

La fórmula para efectuar un cambio de variables en una integral doble tiene una extensión directa para las integrales n-múltiples. Introducimos unas nuevas variables $u_1, \ldots, u_n$ ligadas a $x_1, \ldots, x_n$ por n ecuaciones de la forma

$$x_1 = X_1(u_1, \ldots, u_n), \qquad \ldots, \qquad x_n = X_n(u_1, \ldots, u_n).$$

Pongamos $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{u} = (u_1, \ldots, u_n)$, y $\boldsymbol{X} = (X_1, \ldots, X_n)$. Entonces esas ecuaciones definen una aplicación vectorial

$$\boldsymbol{X}: T \to S$$

de un conjunto T del n-espacio en otro conjunto S del mismo. Supongamos que la aplicación $\boldsymbol{X}$ es uno a uno y derivable con continuidad en T. La fórmula de transformación para las integrales n-múltiples toma la forma

$$\int_S f(\boldsymbol{x})\, d\boldsymbol{x} = \int_T f[\boldsymbol{X}(\boldsymbol{u})]\, |\det D\boldsymbol{X}(\boldsymbol{u})|\, d\boldsymbol{u}, \tag{11.47}$$

en donde $D\boldsymbol{X}(\boldsymbol{u}) = [D_j X_k(\boldsymbol{u})]$ es la matriz jacobiana del campo vectorial $\boldsymbol{X}$. En función de los componentes tenemos

$$D\boldsymbol{X}(\boldsymbol{u}) = \begin{bmatrix} D_1X_1(\boldsymbol{u}) & D_2X_1(\boldsymbol{u}) & \cdots & D_nX_1(\boldsymbol{u}) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ D_1X_n(\boldsymbol{u}) & D_2X_n(\boldsymbol{u}) & \cdots & D_nX_n(\boldsymbol{u}) \end{bmatrix}.$$

Como en el caso bi-dimensional, la fórmula de transformación es válida si $\boldsymbol{X}$ es uno a uno en T y si el determinante jacobiano $J(\boldsymbol{u}) = \det D\boldsymbol{X}(\boldsymbol{u})$ nunca es cero en T. Es también válida si la aplicación deja de ser uno a uno en un subconjunto de T que tenga contenido n-dimensional nulo, o si el determinante jacobiano se anula en un tal subconjunto.

Para el caso tri-dimensional escribimos (x, y, z) en lugar de (x_1, x_2, x_3), (u, v, w) en lugar de (u_1, u_2, u_3) y (X, Y, Z) en lugar de (X_1, X_2, X_3). La fórmula de transformación para las integrales triples toma la forma

$$\iiint_S f(x, y, z)\, dx\, dy\, dz$$

$$= \iiint_T f[X(u, v, w), Y(u, v, w), Z(u, v, w)]\, |J(u, v, w)|\, du\, dv\, dw, \tag{11.48}$$

donde $J(u, v, w)$ es el determinante jacobiano,

$$J(u, v, w) = \begin{vmatrix} \dfrac{\partial X}{\partial u} & \dfrac{\partial Y}{\partial u} & \dfrac{\partial Z}{\partial u} \\[2ex] \dfrac{\partial X}{\partial v} & \dfrac{\partial Y}{\partial v} & \dfrac{\partial Z}{\partial v} \\[2ex] \dfrac{\partial X}{\partial w} & \dfrac{\partial Y}{\partial w} & \dfrac{\partial Z}{\partial w} \end{vmatrix}$$

En el espacio tri-dimensional el determinante jacobiano puede imaginarse como un factor de ampliación de los volúmenes. En efecto, si introducimos la función vectorial $\boldsymbol{r}$ definida por la ecuación

$$\boldsymbol{r}(u, v, w) = X(u, v, w)\boldsymbol{i} + Y(u, v, w)\boldsymbol{j} + Z(u, v, w)\boldsymbol{k},$$

y los vectores

$$V_1 = \frac{\partial \boldsymbol{r}}{\partial u} = \frac{\partial X}{\partial u}\boldsymbol{i} + \frac{\partial Y}{\partial u}\boldsymbol{j} + \frac{\partial Z}{\partial u}\boldsymbol{k},$$

$$V_2 = \frac{\partial \boldsymbol{r}}{\partial v} = \frac{\partial X}{\partial v}\boldsymbol{i} + \frac{\partial Y}{\partial v}\boldsymbol{j} + \frac{\partial Z}{\partial v}\boldsymbol{k},$$

$$V_3 = \frac{\partial \boldsymbol{r}}{\partial w} = \frac{\partial X}{\partial w}\boldsymbol{i} + \frac{\partial Y}{\partial w}\boldsymbol{j} + \frac{\partial Z}{\partial w}\boldsymbol{k},$$

un razonamiento parecido al que se dio en la sección 11.26 sugiere que un ortoedro de dimensiones Δu, Δv, Δw del espacio uvw se transforma en un sólido, que es casi un «paralelepípedo» curvilíneo, en el espacio xyz determinado por los tres vectores $V_1\Delta u$, $V_2\Delta v$, y $V_3\Delta w$. (Ver figura 11.30.) Las caras que limitan este sólido son superficies obtenidas haciendo $u =$ constante, $v =$ constante, y $w =$ constante, respectivamente. El volumen de un paralelepípedo es igual al valor absoluto del producto mixto de los tres vectores que lo determinan, de manera que el volumen del paralelepípedo curvilíneo es aproximadamente igual a

$$|(V_1\,\Delta u)\cdot(V_2\,\Delta v)\times(V_3\,\Delta w)| = |V_1\cdot V_2\times V_3|\,\Delta u\,\Delta v\,\Delta w = |J(u, v, w)|\,\Delta u\,\Delta v\,\Delta w.$$

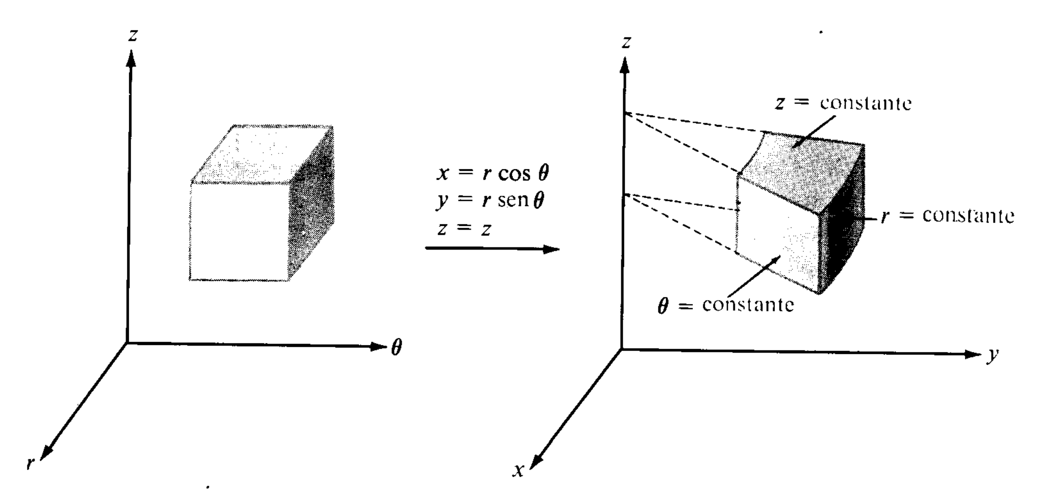


FIGURA 11.30 *Transformación mediante coordenadas cilíndricas.*

11.33 Ejemplos resueltos

En los dos ejemplos que siguen se estudian dos casos particulares importantes de (11.48).

EJEMPLO 1. *Coordenadas cilíndricas.* Ponemos aquí r, θ, z en lugar de u, v, w y definimos la aplicación mediante las ecuaciones

(11.49) $$x = r\cos\theta, \qquad y = r\,\mathrm{sen}\,\theta, \qquad z = z.$$

Es decir, reemplazamos x e y por sus coordenadas polares en el plano xy y no cambiamos la z. También aquí, para conseguir una aplicación uno a uno debemos mantener $r > 0$ y limitar θ a un intervalo de la forma $\theta_0 \leq \theta < \theta_0 + 2\pi$. La figura 11.30 muestra la transformación de un ortoedro del espacio $r\theta z$.

El jacobiano de la aplicación (11.49) es:

$$J(r, \theta, z) = \begin{vmatrix} \cos\theta & \mathrm{sen}\,\theta & 0 \\ -r\,\mathrm{sen}\,\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r(\cos^2\theta + \mathrm{sen}^2\,\theta) = r,$$

y por lo tanto, la fórmula de transformación (11.48) se convierte en

$$\iiint_S f(x, y, z)\,dx\,dy\,dz = \iiint_T f(r\cos\theta, r\,\mathrm{sen}\,\theta, z)r\,dr\,d\theta\,dz.$$

El jacobiano se anula cuando $r = 0$, pero esto no afecta la validez de la fórmula de transformación porque el conjunto de puntos con $r = 0$ tiene medida tri-dimensional 0.

EJEMPLO 2. *Coordenadas esféricas.* En este caso, en lugar de u, v, w ponemos ρ, θ, φ y la aplicación se define por las ecuaciones

$$x = \rho\cos\theta\,\mathrm{sen}\,\varphi, \qquad y = \rho\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\varphi, \qquad z = \rho\cos\varphi.$$

El significado geométrico de ρ, θ, y φ se pone de manifiesto en la figura 11.31. Al objeto de conseguir un aplicación uno a uno limitamos las coordenadas por $\rho > 0$, $0 \leq \theta < 2\pi$, y $0 \leq \varphi < \pi$. Las superficies $\rho =$ constante son esferas concéntricas de centro en el origen, las superficies $\theta =$ constante son planos que pasan por el eje x, y las superficies $\varphi =$ constante son conos circulares cuyo eje es el eje z. Por consiguiente, una caja rectangular en el espacio $\rho\theta\varphi$ se transforma en un sólido del tipo que se ve en la figura 11.31.

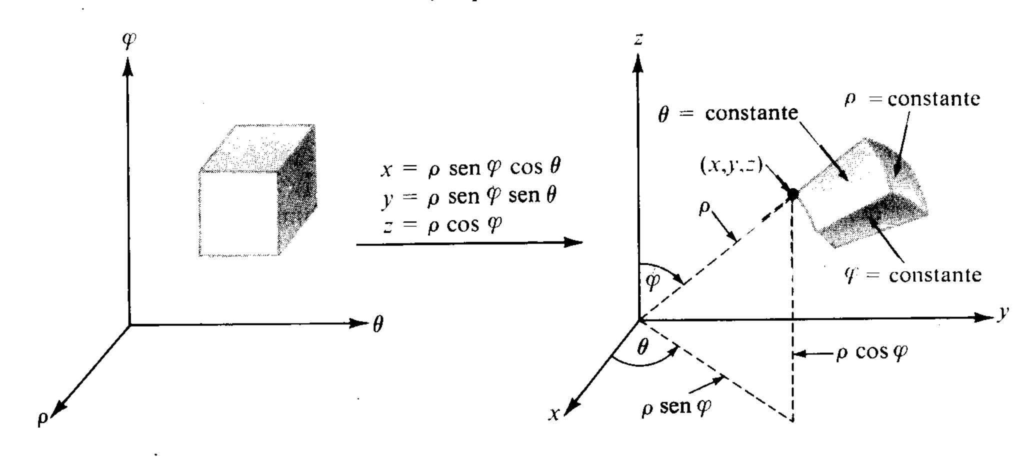


FIGURE 11.31 *Transformación mediante coordenadas esféricas.*

El jacobiano de la aplicación es

$$J(\rho, \theta, \varphi) = \begin{vmatrix} \cos\theta \operatorname{sen}\varphi & \operatorname{sen}\theta \operatorname{sen}\varphi & \cos\varphi \\ -\rho \operatorname{sen}\theta \operatorname{sen}\varphi & \rho\cos\theta \operatorname{sen}\varphi & 0 \\ \rho\cos\theta\cos\varphi & \rho \operatorname{sen}\theta\cos\varphi & -\rho \operatorname{sen}\varphi \end{vmatrix} = -\rho^2 \operatorname{sen}\varphi .$$

Ya que sen $\varphi \leq 0$ si $0 \leq \varphi < \pi$, tenemos $|J(\rho, \theta, \varphi)| = \rho^2$ sen φ y la fórmula de transformación de integrales triples es:

$$\iiint_S f(x, y, z)\, dx\, dy\, dz = \iiint_T F(\rho, \theta, \varphi)\rho^2 \operatorname{sen}\varphi\, d\rho\, d\theta\, d\varphi ,$$

donde $F(\rho, \theta\ \varphi) = f(\rho\cos\theta \operatorname{sen}\varphi, \rho \operatorname{sen}\theta \operatorname{sen}\varphi, \rho\cos\varphi)$. Si bien el jacobiano se anula cuando $\varphi = 0$ la fórmula de transformación es aún válida porque el conjunto de puntos con $\varphi = 0$ tiene medida tri-dimensional 0.

En concepto de volumen puede extenderse a ciertas clases de conjuntos (llamados conjuntos medibles) del *n*-espacio de tal manera que si *S* es medible, su volumen es igual a la integral de la función constante igual a 1 en S. Esto es, si $v(S)$ representa el volumen de *S*, tenemos

$$v(S) = \int \cdots \int_S dx_1 \cdots dx_n .$$

No vamos a intentar describir la clase de los conjuntos para los que es válida

esta fórmula. En lugar de ello, expondremos como puede calcularse la integral en algunos casos particulares.

EJEMPLO 3. *Volumen de un intervalo n-dimensional.* Si S es un intervalo n-dimensional, por ejemplo $S = [a_1, b_1] \times \ldots \times [a_n, b_n]$, la integral múltiple para el cálculo de $v(S)$ es el producto de n integrales uni-dimensionales,

$$v(S) = \int_{a_1}^{b_1} dx_1 \cdots \int_{a_n}^{b_n} dx_n = (b_1 - a_1) \cdots (b_n - a_n).$$

Esto coincide con la fórmula dada antes para el volumen de un intervalo n-dimensional.

EJEMPLO 4. *Volumen de una esfera n-dimensional.* Designemos con $S_n(a)$ la esfera n-dimensional (o n-esfera) de radio a dada por

$$S_n(a) = \{(x_1, \ldots, x_n) \mid x_1^2 + \cdots + x_n^2 \leq a^2\},$$

y sea

$$V_n(a) = \int \cdots \int_{S_n(a)} dx_1 \cdots dx_n,$$

el volumen de $S_n(a)$. Demostraremos que

(11.50) $$V_n(a) = \frac{\pi^{n/2}}{\Gamma(\frac{1}{2}n + 1)} a^n,$$

donde Γ es la función gamma. Para $n = 1$ la fórmula da $V_1(a) = 2a$, que es la longitud del intervalo $[-a, a]$. Para $n = 2$ da $V_2(a) = \pi a^2$, que es el área de un disco circular de radio a. Demostraremos (11.50) para $n \geqslant 3$.

Primero demostraremos que para todo $a > 0$ tenemos

(11.51) $$V_n(a) = a^n V_n(1).$$

Dicho de otro modo, el volumen de una esfera de radio a es a^n multiplicado por el volumen de una esfera de radio 1. Para probarlo hacemos el cambio lineal de variable $\boldsymbol{x} = a\boldsymbol{u}$ para aplicar $S_n(1)$ en $S_n(a)$. La aplicación tiene jacobiano igual a a^n. Luego

$$V_n(a) = \int \cdots \int_{S_n(a)} dx_1 \cdots dx_n = \int \cdots \int_{S_n(1)} a^n \, du_1 \cdots du_n = a^n V_n(1).$$

Esto demuestra (11.51). Por consiguiente, para demostrar (11.50) basta probar que

(11.52) $$V_n(1) = \frac{\pi^{n/2}}{\Gamma(\frac{1}{2}n + 1)}.$$

Observemos primero que $x_1^2 + \cdots + x_n^2 \leq 1$ si y sólo si

$$x_1^2 + \cdots + x_{n-2}^2 \leq 1 - x_{n-1}^2 - x_n^2 \qquad \text{y} \qquad x_{n-1}^2 + x_n^2 \leq 1.$$

Por lo tanto podemos escribir la integral que da $V_n(1)$ como la iteración de una integral $(n - 2)$-múltiple y una integral doble, del modo siguiente:

(11.53) $$V_n(1) = \iint\limits_{x_{n-1}^2+x_n^2\leq 1} \left[\quad \int\cdots\int\limits_{x_1^2+\cdots+x_{n-2}^2\leq 1-x_{n-1}^2-x_n^2} dx_1 \cdots dx_{n-2}\right] dx_{n-1}\, dx_n.$$

La integral interior está extendida a la esfera $(n - 2)$-dimensional $S_{n-2}(R)$, siendo $R = \sqrt{1 - x_{n-1}^2 - x_n^2}$, con lo que es igual a

$$V_{n-2}(R) = R^{n-2}V_{n-2}(1) = (1 - x_{n-1}^2 - x_n^2)^{n/2-1}\, V_{n-2}(1).$$

Ponemos ahora x en lugar de x_{n-1} e y en lugar de x_n. Entonces (11.53) toma la forma

$$V_n(1) = V_{n-2}(1) \iint\limits_{x^2+y^2\leq 1} (1 - x^2 - y^2)^{n/2-1}\, dx\, dy.$$

La integral doble la calculamos transformándola a coordenadas polares y obtenemos

$$V_n(1) = V_{n-2}(1) \int_0^{2\pi} \int_0^1 (1 - r^2)^{n/2-1}\, r\, dr\, d\theta = V_{n-2}(1)\, \frac{2\pi}{n}.$$

Es decir, los números $V_n(1)$ satisfacen la fórmula de recurrencia

$$V_n(1) = \frac{2\pi}{n}\, V_{n-2}(1) \qquad \text{si } n \geq 3.$$

Pero la sucesión de números $\{f(n)\}$ definida por

$$f(n) = \frac{\pi^{n/2}}{\Gamma(\frac{1}{2}n + 1)}$$

satisface la misma fórmula de recurrencia debido a que $\Gamma(s + 1) = s\Gamma(s)$. Asi-

mismo, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ (véase ejercicio 16, sección 11.28), con lo que $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$ y $f(1) = V_1(1) = 2$. También, $f(2) = V_2(1) = \pi$, luego tenemos $f(n) = V_n(1)$ para todo $n \geqslant 1$. Esto demuestra (11.52).

11.34 Ejercicios

Calcular cada una de las integrales triples en los ejercicios del 1 al 5. Representar en cada caso la región de integración. Supóngase la existencia de todas las integrales que se citan.

1. $\iiint_S xy^2z^3\,dx\,dy\,dz$, siendo S el sólido limitado por la superficie $z = xy$ y los planos $y = x$, $x = 1$, y $z = 0$.
2. $\iiint_S (1 + x + y + z)^{-3}\,dx\,dy\,dz$, siendo S el sólido limitado por los tres planos coordenados y el plano $x + y + z = 1$.
3. $\iiint_S xyz\,dx\,dy\,dz$, siendo $S = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1, x \geq 0, y \geq 0, z \geq 0\}$.
4. $\iiint_S \left(\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2}\right) dx\,dy\,dz$, siendo S el sólido limitado por el elipsoide $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$.
5. $\iiint_S \sqrt{x^2 + y^2}\,dx\,dy\,dz$, siendo S el sólido formado por la hoja superior del cono $z^2 = x^2 + y^2$ y el plano $z = 1$.

En los ejercicios 6, 7, y 8, una integral triple $\iiint_S f(x, y, z)\,dx\,dy\,dz$ de una función positiva se reduce a la integral reiterada dada. En cada caso dibujar la región de integración S mostrando su proyección sobre el plano xy. Expresar entonces la integral triple como una o más integrales reiteradas en las que la primera integración se efectúa respecto a y.

6. $\displaystyle\int_0^1\left(\int_0^{1-x}\left[\int_0^{x+y} f(x, y, z)\,dz\right]dy\right)dx.$
7. $\displaystyle\int_{-1}^1\left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\left[\int_{\sqrt{x^2-y^2}}^1 f(x, y, z)\,dz\right]dy\right)dx.$
8. $\displaystyle\int_0^1\left(\int_0^1\left[\int_0^{x^2+y^2} f(x, y, z)\,dz\right]dy\right)dx.$
9. Demostrar que:
$$\int_0^x\left(\int_0^v\left[\int_0^u f(t)\,dt\right]du\right)dv = \frac{1}{2}\int_0^x (x - t)^2 f(t)\,dt.$$

Calcular las integrales de los ejercicios 10, 11 y 12 pasando a coordenadas cilíndricas. Puede suponerse la existencia de todas las integrales que se citan.

10. $\iiint_S (x^2 + y^2)\,dx\,dy\,dz$, siendo S el sólido limitado por la superficie $x^2 + y^2 = 2z$ y el plano $z = 2$.

11. $\iiint_S dx\,dy\,dz$, siendo S el sólido limitado por los tres planos coordenados, la superficie $z = x^2 + y^2$, y el plano $x + y = 1$.

12. $\iiint_S (y^2 + z^2)\,dx\,dy\,dz$, siendo S un cono recto de revolución, de altura h, base situada en el plano xy y de radio a, y eje en el eje z.

Calcular las integrales 13, 14, y 15 mediante la transformación a coordenadas esféricas.

13. $\iiint_S dx\,dy\,dz$, siendo S una esfera de radio a y centro en el origen.

14. $\iiint_S dx\,dy\,dz$, siendo S el sólido limitado entre dos esferas concéntricas de radios a y b, $(0 < a < b)$ y con centro en el origen.

15. $\iiint_S [(x - a)^2 + (y - b)^2 + (z - c)^2]^{-1/2}\,dx\,dy\,dz$, siendo S una esfera de radio R y centro en el origen, y (a, b, c) es un punto fijo en el exterior de esa esfera.

16. Se pueden definir unas coordenadas esféricas generalizadas mediante la aplicación siguiente:

$$x = a\rho \cos^m \theta \operatorname{sen}^n \varphi, \qquad y = b\rho \operatorname{sen}^m \theta \operatorname{sen}^n \varphi, \qquad z = c\rho \cos^n \varphi,$$

donde a, b, c, m y n son constantes positivas. Demostrar que el jacobiano es igual a

$$-abcmn\rho^2 \cos^{m-1} \theta \operatorname{sen}^{m-1} \theta \cos^{n-1} \varphi \operatorname{sen}^{2n-1} \varphi.$$

Las integrales triples pueden emplearse para calcular volúmenes, masas, centros de gravedad, momentos de inercia, y otros conceptos físicos asociados a sólidos. Si S es un sólido su volumen V viene dado por la integral triple

$$V = \iiint_S dx\,dy\,dz.$$

Si al sólido se asigna una densidad $f(x, y, z)$ en cada uno de sus puntos (x, y, z) (masa por unidad de volumen), su masa M es:

$$M = \iiint_S f(x, y, z)\,dx\,dy\,dz,$$

y su centro de gravedad el punto $(\bar{x}, \bar{y}, \bar{z})$ con coordenadas

$$\bar{x} = \frac{1}{M} \iiint_S x f(x, y, z)\,dx\,dy\,dz,$$

y con cálculo parecido $\bar{y}$ y $\bar{z}$. El momento de inercia I_{xy} respecto al plano xy se define por la igualdad

$$I_{xy} = \iiint_S z^2 f(x, y, z)\,dx\,dy\,dz$$

y fórmulas parecidas para I_{yz} e I_{zx}. El momento de inercia I_L respecto a una recta L se define como

$$I_L = \iiint_S \delta^2(x, y, z) f(x, y, z)\, dx\, dy\, dz,$$

donde $\delta(x, y, z)$ representa la distancia de un punto genérico de S a la recta L.

17. Demostrar que los momentos de inercia respecto a los ejes coordenados son

$$I_x = I_{xy} + I_{xz}, \qquad I_y = I_{yx} + I_{yz}, \qquad I_z = I_{zx} + I_{zy}.$$

18. Calcular el volumen de un sólido limitado por encima por la esfera $x^2 + y^2 + z^2 = 5$ y por debajo por el paraboloide $x^2 + y^2 = 4z$.
19. Calcular el volumen de un sólido limitado por el plano xy, el cilindro $x^2 + y^2 = 2x$, y el cono $z = \sqrt{x^2 + y^2}$.
20. Calcular la masa del sólido limitado por dos esferas concéntricas de radios a y b, $(0 < a < b)$, si la densidad en cada punto es igual al cuadrado de su distancia al centro.
21. Un cono circular recto homogéneo tiene altura h. Demostrar que la distancia de su centroide a la base es $h/4$.
22. Determinar el centro de gravedad de un cono de altura h si su densidad en cada punto es proporcional a la distancia de ese punto a la base.
23. Ídem, suponiendo que la densidad en cada punto es proporcional a la distancia de ese punto al eje del cono.
24. Un sólido está limitado por dos hemisferios concéntricos de radios a y b, siendo $0 < a < b$. Si la densidad es constante, encontrar el centro de gravedad.
25. Determinar el centro de gravedad de un cubo de lado h si su densidad en cada punto es proporcional al cuadrado de la distancia de ese punto a un vértice de la base. (Tomar la base en el plano xy y situar las aristas en los ejes coordenados.)
26. Un cono recto de revolución tiene altura h, radio en la base a, densidad constante, y masa M. Encontrar su momento de inercia respecto a un eje paralelo a la base y que pasa por el vértice.
27. Calcular el momento de inercia de una esfera de radio R y masa M respecto a un diámetro si la densidad es constante.
28. Calcular el momento de inercia de un cilindro de radio a y masa M si su densidad en cada punto es proporcional a la distancia de ese punto al eje del cilindro.
29. El tallo de una seta es un cilindro recto de revolución de diámetro 1 y longitud 2, y su cabeza es un hemisferio de radio R. Si la seta es un sólido homogéneo con simetría axial, y su centro de gravedad está situado en el plano en el que el tallo se une a la cabeza, calcular R.
30. Un satélite artificial tiene una cubierta compuesta de porciones de dos cilindros de revolución de igual diámetro D cuyos ejes se cortan en ángulo recto. Hay el propósito de transportar el satélite a Cabo Kennedy en una caja cúbica de interior D. Demostrar que una tercera parte del volumen de la caja se pierde.
31. Designemos con $S_n(a)$ el siguiente conjunto del n-espacio, siendo $a > 0$:

$$S_n(a) = \{(x_1, \ldots, x_n) \mid |x_1| + \cdots + |x_n| \le a\}.$$

Cuando $n = 2$ el conjunto es un cuadrado con vértices en $(0, \pm a)$ y $(\pm a, 0)$. Cuando

$n = 3$ es un octaedro con vértices en $(0, 0, \pm a)$, $(0, \pm a, 0)$, y $(\pm a, 0, 0)$. Llamemos $V_n(a)$ al volumen de $S_n(a)$ dado por

$$V_n(a) = \int_{S_n(a)} \cdots \int dx_1 \cdots dx_n .$$

a) Demostrar que $V_n(a) = a^n V_n(1)$.

b) Para $n \geq 2$, expresar la integral que da $V_n(1)$ como una iteración de una integral uni-dimensional y una integral $(n - 1)$-múltiple y demostrar que

$$V_n(1) = V_{n-1}(1) \int_{-1}^{1} (1 - |x|)^{n-1}\, dx = \frac{2}{n} V_{n-1}(1) .$$

c) Hacer uso de los apartados a) y b) para deducir que $V_n(a) = \dfrac{2^n a^n}{n!}$

32. Designemos con $S_n(a)$ el siguiente conjunto del n-espacio, siendo $a > 0$ y $n \geq 2$:

$$S_n(a) = \{(x_1, \ldots, x_n) \mid |x_i| + |x_n| \leq a \text{ para cada } i = 1, \ldots, n - 1\} .$$

a) Esbozar un gráfico de $S_n(1)$ cuando $n = 2$ y cuando $n = 3$.

b) Póngase $V_n(a) = \int_{S_n(a)} \cdots \int dx_1 \cdots dx_n$, y demostrar que $V_n(a) = a^n V_n(1)$.

c) Expresar la integral que da $V_n(1)$ como una iteración de una integral uni-dimensional y una integral $(n - 1)$-múltiple y deducir que $V_n(a) = 2^n a^n / n$.

33. a) En relación con el ejemplo 4, pág. 502. Expresar la integral que da $V_n(1)$, (volumen de la esfera n-dimensional unidad) como la iteración de una integral $(n - 1)$-múltiple y una integral uni-dimensional y con ello demostrar que

$$V_n(1) = 2V_{n-1}(1) \int_0^1 (1 - x^2)^{(n-1)/2}\, dx.$$

b) Con el apartado a) y la ecuación (11.52) deducir que

$$\int_0^{\pi/2} \cos^n t\, dt = \frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2} + 1\right)} .$$

12

INTEGRALES DE SUPERFICIE

12.1 Representación paramétrica de una superficie

Este capítulo trata de las integrales de superficie y sus aplicaciones. Puede imaginarse la integral de superficie como el equivalente en dos dimensiones a una integral de línea siendo la región de integración una superficie en lugar de una curva. Antes de estudiar las integrales de superficie, tenemos que ponernos de acuerdo en lo que es una superficie.

Hablando sin precisión, una superficie es el lugar de un punto que se mueve en el espacio con dos grados de libertad. En la parte de Geometría analítica del Volumen I vimos dos métodos para expresar analíticamente un tal lugar. Uno es la *representación implícita* en el que se considera una superficie como un conjunto de puntos (x, y, z) que satisfacen una ecuación de la forma $F(x, y, z) = 0$. Algunas veces podemos despejar en la ecuación una de las coordenadas en función de las otras dos, por ejemplo z en función de x e y. Cuando eso es posible obtenemos una *representación explícita* dada por una o varias ecuaciones de la forma $z = f(x, y)$. Por ejemplo, una esfera de radio 1 y centro en el origen tiene la representación implícita $x^2 + y^2 + z^2 - 1 = 0$. Al despejar z se obtienen dos soluciones, $z = \sqrt{1 - x^2 - y^2}$ y $z = -\sqrt{1 - x^2 - y^2}$. La primera es la representación explícita de la semiesfera superior y la segunda la de la inferior.

Existe un tercer método de representación de superficies que es más útil en el estudio de las mismas; es la representación *paramétrica* o *vectorial* por medio de tres ecuaciones que expresan x, y, z en función de dos parámetros u y v:

$$(12.1) \qquad x = X(u, v), \qquad y = Y(u, v), \qquad z = Z(u, v).$$

Aquí el punto (u, v) puede variar en un conjunto conexo bidimensional T en el plano uv, y los puntos (x, y, z) correspondientes constituyen una porción de super-

ficie en el espacio xyz. Este método es análogo al de la representación de una curva en E_3 mediante tres ecuaciones con un parámetro. La presencia de los dos parámetros en (12.1) permite transmitir dos grados de libertad al punto (x, y, z), como sugiere la figura 12.1. Otro modo de expresar la misma idea consiste en decir que una superficie es la imagen de una región plana T por medio de la aplicación definida por (12.1).

Si introducimos el radio vector $\boldsymbol{r}$ que une el origen a un punto genérico (x, y, z) de la superficie, podemos combinar las tres ecuaciones paramétricas (12.1)

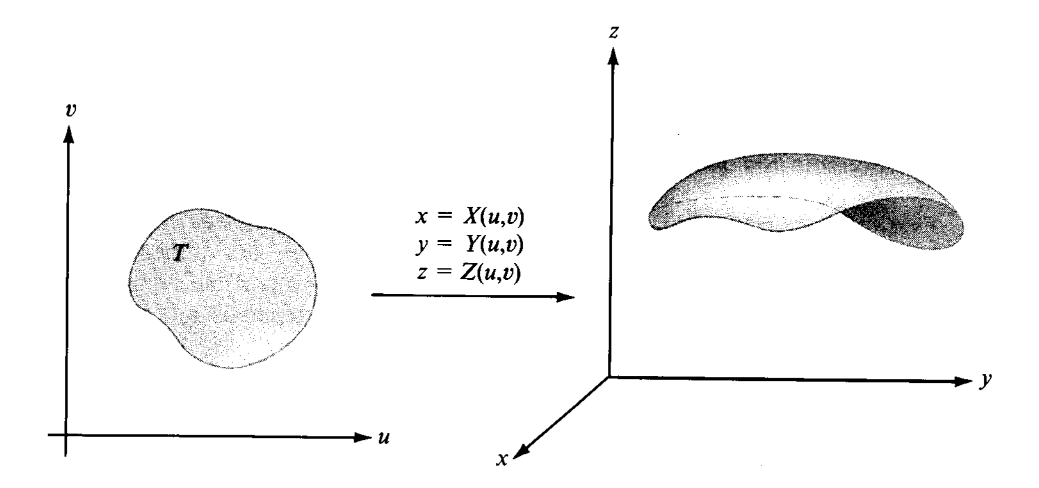


FIGURA 12.1 *Representación paramétrica de una superficie.*

en una ecuación vectorial de la forma

$$(12.2)\qquad \boldsymbol{r}(u, v) = X(u, v)\boldsymbol{i} + Y(u, v)\boldsymbol{j} + Z(u, v)\boldsymbol{k}, \qquad \text{donde} \quad (u, v) \in T.$$

Ésta es la llamada *ecuación vectorial* de la superficie.

Existen, naturalmente, muchas representaciones paramétricas de la misma superficie. Una de ellas puede obtenerse siempre a partir de la forma explícita $z = f(x, y)$ tomando $X(u, v) = u$, $Y(u, v) = v$, $Z(u, v) = f(u, v)$. Por otra parte, si es posible eliminar u y v en las ecuaciones paramétricas —por ejemplo, si podemos resolver las dos primeras ecuaciones (12.1) respecto a u y v en función de x e y y sustituimos en la tercera— obtenemos la representación explícita $z = f(x, y)$.

EJEMPLO 1. *Representación paramétrica de una esfera.* Las tres ecuaciones

$$(12.3)\qquad x = a\cos u\cos v, \qquad y = a\operatorname{sen} u\cos v, \qquad z = a\operatorname{sen} v$$

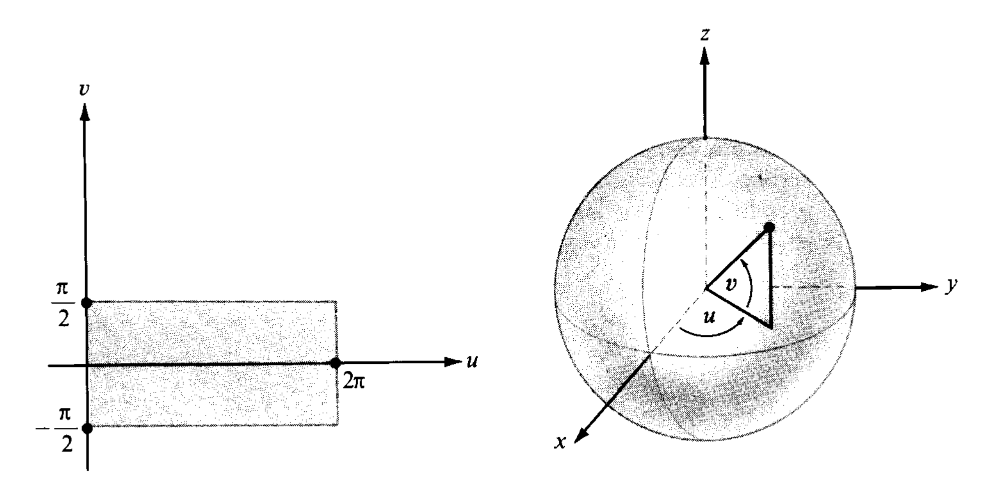


FIGURA 12.2 *Representación paramétrica de una esfera.*

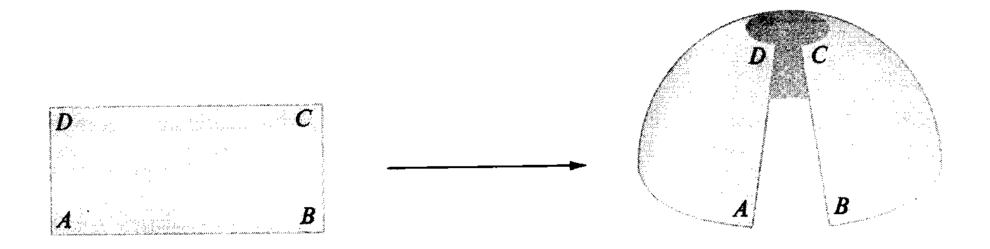


FIGURA 12.3 *Deformación de un rectángulo en un hemisferio.*

representan una esfera de radio a y centro en el origen. Si elevamos al cuadrado las tres ecuaciones (12.3) y sumamos resulta $x^2 + y^2 + z^2 = a^2$, y vemos que todo punto (x, y, z) que satisface (12.3) está en la esfera. Los parámetros u y v en este ejemplo pueden interpretarse geométricamente como los ángulos dibujados en la figura 12.2. Si hacemos que el punto (u, v) varíe en el rectángulo $T = [0, 2\pi] \times [-\frac{1}{2}\pi, \frac{1}{2}\pi]$, los puntos determinados por (12.3) describen toda la esfera. El hemisferio superior es la imagen del rectángulo $[0, 2\pi] \times [0, \frac{1}{2}\pi]$ y el inferior es imagen de $[0, 2\pi] \times [-\frac{1}{2}\pi, 0]$. La figura 12.3 nos da una idea de como el rectángulo $[0, 2\pi] \times [0, \frac{1}{2}\pi]$ es aplicado en el hemisferio superior. Imaginemos que el rectángulo es de un material plástico flexible capaz de estirarse o encogerse. La figura 12.3 muestra el rectángulo convertido por deformación en un hemisferio. La base AB se transforma en el ecuador, los lados opuestos AD y BC se hacen coincidir, y el lado superior DC degenera en un punto (el polo Norte).

EJEMPLO 2. *Representación paramétrica de un cono.* La ecuación vectorial

$$\boldsymbol{r}(u, v) = v \operatorname{sen} \alpha \cos u\, \boldsymbol{i} + v \operatorname{sen} \alpha \operatorname{sen} u\, \boldsymbol{j} + v \cos \alpha\, \boldsymbol{k}$$

representa el cono circular recto dibujado en la figura 12.4, donde α es el semiángulo en el vértice. También aquí, los parámetros u, v, admiten interpretaciones geométricas; v es la distancia desde el vértice al punto (x, y, z) del cono, y u es el ángulo polar. Cuando (u, v) varía en un rectángulo de la forma $[0, 2\pi] \times$

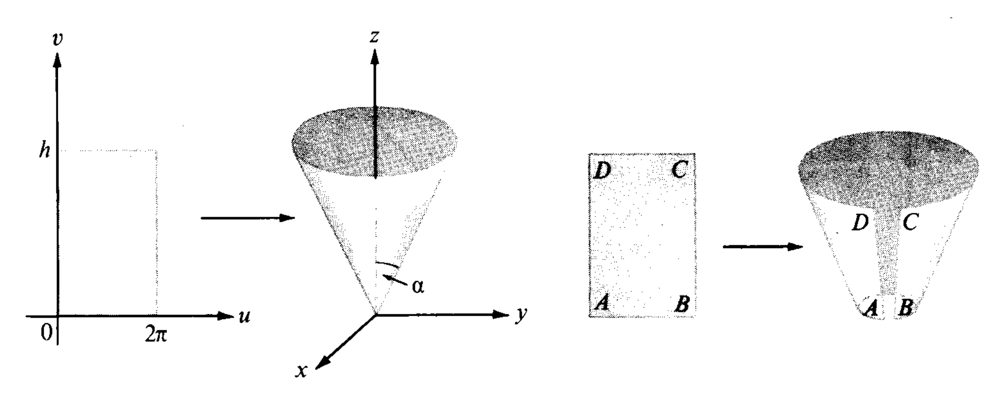


FIGURA 12.4 *Representación paramétrica de un cono.*

FIGURA 12.5 *Deformación de un rectángulo en un cono.*

$[0, h]$, los correspondientes puntos (x, y, z) describen un cono de altura $h \cos \alpha$. Un rectángulo plástico puede, por deformación, convertirse en un cono haciendo coincidir los lados AD y BC, como indica la figura 12.5, y convirtiendo el lado AB en un punto (el vértice del cono). La superficie de la figura 12.5 muestra un estado intermedio de la deformación.

En el estudio general de las superficies, las funciones X, Y, y Z que aparecen en las ecuaciones paramétricas (12.1) o en la ecuación vectorial (12.2) se suponen continuas en T. La imagen de T a través de la aplicación $\boldsymbol{r}$ se llama *superficie paramétrica* y se representará mediante el símbolo $\boldsymbol{r}(T)$. En muchos de los ejemplos que comentaremos, T será un rectángulo, un disco circular, o algún otro conjunto conexo simple limitado por una curva cerrada simple. Si la función $\boldsymbol{r}$ es uno a uno en T, la imagen $\boldsymbol{r}(T)$ de denominará *superficie paramétrica simple.* En tal caso, puntos distintos de T se aplican en puntos distintos de la superficie. En particular, toda curva cerrada simple en T se aplica en una curva cerrada simple situada en la superficie.

Una superficie paramétrica $\boldsymbol{r}(T)$ puede degenerar en un punto o en una curva. Por ejemplo, si las tres funciones X, Y, Z son constantes, la imagen $\boldsymbol{r}(T)$ es un solo punto. Si X, Y y Z son independientes de v, la imagen $\boldsymbol{r}(T)$ es una curva.

Otro ejemplo de superficie degenerada se presenta cuando $X(u, v) = u + v$, $Y(u, v) = (u + v)^2$, y $Z(u, v) = (u + v)^3$, siendo $T = [0, 1] \times [0, 1]$. Si escribimos $t = u + v$ vemos que la superficie degenera en una curva alabeada de ecuaciones paramétricas $x = t$, $y = t^2$ y $z = t^3$, donde $0 \leq t \leq 2$. Estos casos excepcionales pueden evitarse imponiendo ciertas restricciones a la función $\boldsymbol{r}$ que define la aplicación, como se explica en la siguiente sección.

12.2 Producto vectorial fundamental

Consideremos una superficie representada por la ecuación vectorial

$$\boldsymbol{r}(u, v) = X(u, v)\boldsymbol{i} + Y(u, v)\boldsymbol{j} + Z(u, v)\boldsymbol{k}, \qquad \text{donde} \quad (u, v) \in T.$$

Si X, Y, y Z son derivables en T podemos considerar los dos vectores

$$\frac{\partial \boldsymbol{r}}{\partial u} = \frac{\partial X}{\partial u}\boldsymbol{i} + \frac{\partial Y}{\partial u}\boldsymbol{j} + \frac{\partial Z}{\partial u}\boldsymbol{k}$$

y

$$\frac{\partial \boldsymbol{r}}{\partial v} = \frac{\partial X}{\partial v}\boldsymbol{i} + \frac{\partial Y}{\partial v}\boldsymbol{j} + \frac{\partial Z}{\partial v}\boldsymbol{k}.$$

El producto vectorial $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ se denominará *producto vectorial fundamental* de la representación $\boldsymbol{r}$. Sus componentes pueden expresarse como determinantes jacobianos. En efecto, tenemos,

$$(12.4) \quad \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial X}{\partial u} & \frac{\partial Y}{\partial u} & \frac{\partial Z}{\partial u} \\ \frac{\partial X}{\partial v} & \frac{\partial Y}{\partial v} & \frac{\partial Z}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial Y}{\partial u} & \frac{\partial Z}{\partial u} \\ \frac{\partial Y}{\partial v} & \frac{\partial Z}{\partial v} \end{vmatrix} \boldsymbol{i} + \begin{vmatrix} \frac{\partial Z}{\partial u} & \frac{\partial X}{\partial u} \\ \frac{\partial Z}{\partial v} & \frac{\partial X}{\partial v} \end{vmatrix} \boldsymbol{j} + \begin{vmatrix} \frac{\partial X}{\partial u} & \frac{\partial Y}{\partial u} \\ \frac{\partial X}{\partial v} & \frac{\partial Y}{\partial v} \end{vmatrix} \boldsymbol{k}$$

$$= \frac{\partial(Y, Z)}{\partial(u, v)}\boldsymbol{i} + \frac{\partial(Z, X)}{\partial(u, v)}\boldsymbol{j} + \frac{\partial(X, Y)}{\partial(u, v)}\boldsymbol{k}.$$

Si (u, v) es un punto en T en el cual $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ son continuas y el producto vectorial fundamental no es nulo, el punto imagen $\boldsymbol{r}(u, v)$ se llama *punto regular* de $\boldsymbol{r}$. Los puntos en los que no son continuas $\partial \boldsymbol{r}/\partial u$ o $\partial \boldsymbol{r}/\partial v$ o bien $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v = \boldsymbol{O}$ se llaman *puntos singulares* de $\boldsymbol{r}$. Una superficie $\boldsymbol{r}(T)$ se llama *regular* si todos sus puntos son regulares. Toda superficie tiene más de una representación paramétrica. Algunos de los ejemplos que luego se comentan ponen de manifiesto que un punto de una superficie puede ser regular para una represen-

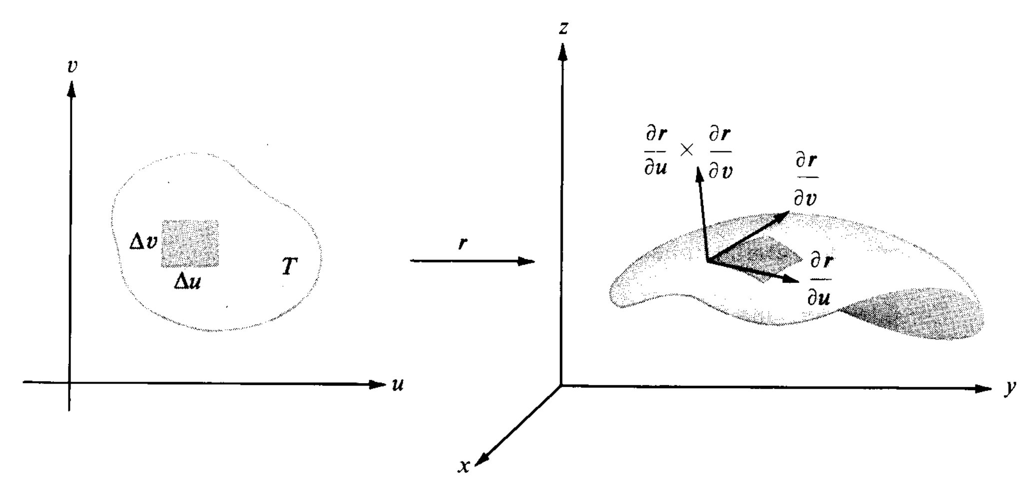


FIGURA 12.6 *Interpretación geométrica de los vectores* $\frac{\partial \boldsymbol{r}}{\partial u}$, $\frac{\partial \boldsymbol{r}}{\partial v}$, *y* $\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}$.

tación y singular para otra. Seguidamente explicamos el significado geométrico de los conceptos de puntos regulares y singulares.

Consideremos en T un segmento rectilíneo horizontal. Su imagen por $\boldsymbol{r}$ es una curva (llamada u-curva) situada en la superficie $\boldsymbol{r}(T)$. Para v fija, imaginemos que el parámetro u represente el tiempo. El vector $\partial \boldsymbol{r}/\partial u$ es el vector velocidad de esta curva. Cuando u se incrementa en Δu, un punto situado al principio en $\boldsymbol{r}(u, v)$ se desplaza a lo largo de una u-curva una distancia aproximadamente igual a $\|\partial \boldsymbol{r}/\partial u\| \Delta u$ puesto que $\|\partial \boldsymbol{r}/\partial u\|$ representa la velocidad a lo largo de la u-curva. Análogamente, para u fija un punto de una v-curva se desplaza en el tiempo Δv una distancia aproximadamente igual a $\|\partial \boldsymbol{r}/\partial v\| \Delta v$. Un rectángulo en T que tenga un área $\Delta u\, \Delta v$ se convierte en una porción de $\boldsymbol{r}(T)$ que aproximaremos por un paralelogramo determinado por los vectores $(\partial \boldsymbol{r}/\partial u)\,\Delta u$ y $(\partial \boldsymbol{r}/\partial v)\,\Delta v$. (Ver figura 12.6.) El área del paralelogramo determinado por $(\partial \boldsymbol{r}/\partial u)\,\Delta u$ y $(\partial \boldsymbol{r}/\partial v)\,\Delta v$ es el módulo de su producto vectorial

$$\left\| \frac{\partial \boldsymbol{r}}{\partial u} \Delta u \times \frac{\partial \boldsymbol{r}}{\partial v} \Delta v \right\| = \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| \Delta u\, \Delta v .$$

Por consiguiente la longitud del producto vectorial fundamental puede imaginarse como un factor de proporcionalidad de las áreas. En los puntos en los que este producto vectorial es nulo el paralelogramo degenera en una curva o en un punto. En cada punto regular los vectores $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ determinan un plano que tiene el vector $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ como normal. En la próxima sección demostraremos que $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ es normal a toda curva regular en la superficie; por esta razón el plano determinado por $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ se llama *plano tangente* a la super-

ficie. La continuidad de $\partial r/\partial u$ y $\partial r/\partial v$ implica la continuidad de $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$; esto. a su vez, significa que el plano tangente se mueve con continuidad en una superficie regular. Así vemos que la continuidad de $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ evita la presencia de aristas o «puntas» en la superficie; la no anulación de $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ evita los casos degenerados antes citados.

EJEMPLO 1. *Superficies con representación explícita, $z = f(x, y)$.* Para una superficie con una representación explícita de la forma $z = f(x, y)$, podemos usar x e y como parámetros, lo que nos da la ecuación vectorial

$$\boldsymbol{r}(x, y) = x\boldsymbol{i} + y\boldsymbol{j} + f(x, y)\boldsymbol{k}.$$

Esta representación nos da siempre una superficie paramétrica simple. La región T se denomina proyección de la superficie sobre el plano xy. (En la figura 12.7 se muestra un ejemplo.) Para calcular el producto vectorial fundamental observemos que

$$\frac{\partial \boldsymbol{r}}{\partial x} = \boldsymbol{i} + \frac{\partial f}{\partial x}\boldsymbol{k} \qquad \text{y} \qquad \frac{\partial \boldsymbol{r}}{\partial y} = \boldsymbol{j} + \frac{\partial f}{\partial y}\boldsymbol{k},$$

si f es diferenciable. Esto nos da

$$(12.5) \qquad \frac{\partial \boldsymbol{r}}{\partial x} \times \frac{\partial \boldsymbol{r}}{\partial y} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 0 & \dfrac{\partial f}{\partial x} \\ 0 & 1 & \dfrac{\partial f}{\partial y} \end{vmatrix} = -\frac{\partial f}{\partial x}\boldsymbol{i} - \frac{\partial f}{\partial y}\boldsymbol{j} + \boldsymbol{k}.$$

Puesto que el componente z de $\partial \boldsymbol{r}/\partial x \times \partial \boldsymbol{r}/\partial y$ es 1, el producto vectorial fundamental nunca es cero. Luego los únicos puntos singulares que pueden presentarse por esta representación son puntos en los que al menos una de las derivadas parciales $\partial f/\partial x$ o $\partial f/\partial y$ no es continua.

Un caso típico es la ecuación $z = \sqrt{1 - x^2 - y^2}$, que representa un hemisferio de radio 1 y centro en el origen, si $x^2 + y^2 \leq 1$. La ecuación vectorial

$$\boldsymbol{r}(x, y) = x\boldsymbol{i} + y\boldsymbol{j} + \sqrt{1 - x^2 - y^2}\boldsymbol{k}$$

aplica el disco unidad $T = \{(x, y) \mid x^2 + y^2 \leq 1\}$ sobre el hemisferio y dicha aplicación es uno a uno. Las derivadas parciales $\partial \boldsymbol{r}/\partial x$ y $\partial \boldsymbol{r}/\partial y$ existen y son conti-

nuas en todo el interior del disco, pero no existen en la periferia. Por consiguiente, todo punto del ecuador es un punto singular de esta representación.

EJEMPLO 2. Consideremos el mismo hemisferio del ejemplo anterior, pero esta vez como imagen del rectángulo $T = [0, 2\pi] \times [0, \frac{1}{2}\pi]$ a través de la aplicación

$$\boldsymbol{r}(u, v) = a \cos u \cos v\, \boldsymbol{i} + a \operatorname{sen} u \cos v\, \boldsymbol{j} + a \operatorname{sen} v\, \boldsymbol{k}.$$

Los vectores $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ vienen dados por las fórmulas

$$\frac{\partial \boldsymbol{r}}{\partial u} = -a \operatorname{sen} u \cos v\, \boldsymbol{i} + a \cos u \cos v\, \boldsymbol{j},$$

$$\frac{\partial \boldsymbol{r}}{\partial v} = -a \cos u \operatorname{sen} v\, \boldsymbol{i} - a \sin u \operatorname{sen} v\, \boldsymbol{j} + a \cos v \boldsymbol{k}.$$

Un sencillo cálculo muestra que su producto vectorial es igual a

$$\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} = a \cos v\, \boldsymbol{r}(u, v).$$

La imagen de T no es una superficie paramétrica simple a causa de que esta aplicación no es uno a uno en T. En efecto, todo punto del segmento rectilíneo $v = \frac{1}{2}\pi$, $0 \leq u \leq 2\pi$ se aplica en el punto $(0, 0, a)$ (el polo Norte). También, por la periodicidad del seno y coseno, $\boldsymbol{r}$ toma los mismos valores en los puntos $(0, v)$ y $(2\pi, v)$, de modo que los lados izquierdo y derecho de T se aplican en la misma curva, que es un arco que une el polo Norte al punto $(a, 0, 0)$ del ecuador. (Véase figura 12.3.) Los vectores $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ son continuos en todo T. Puesto que $\| \partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v \| = a^2 \cos v$, los únicos puntos singulares de esta representación se presentan cuando $\cos v = 0$. No hay otro punto singular que el polo Norte.

12.3 El producto vectorial fundamental, considerado como una normal a la superficie

Consideremos una superficie paramétrica regular $\boldsymbol{r}(T)$, y sea C^* una curva regular en T. La imagen $C = \boldsymbol{r}(C^*)$ es entonces una curva regular situada en la superficie. Demostraremos que en cada punto de C el vector $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ es normal a C, como se ve en la figura 12.6.

Supongamos que C^* está descrita por una función $\boldsymbol{\alpha}$ definida en un intervalo $[a, b]$, por ejemplo sea

$$\boldsymbol{\alpha}(t) = U(t)\boldsymbol{i} + V(t)\boldsymbol{j}.$$

Entonces la imagen de la curva C está representada por la función compuesta

$$\boldsymbol{\rho}(t) = \boldsymbol{r}[\boldsymbol{\alpha}(t)] = X[\boldsymbol{\alpha}(t)]\boldsymbol{i} + Y[\boldsymbol{\alpha}(t)]\boldsymbol{j} + Z[\boldsymbol{\alpha}(t)]\boldsymbol{k}.$$

Queremos demostrar que la derivada $\boldsymbol{\rho}'(t)$ es perpendicular al vector $\partial\boldsymbol{r}/\partial u \times \partial\boldsymbol{r}/\partial v$ cuando las derivadas parciales $\partial\boldsymbol{r}/\partial u$ y $\partial\boldsymbol{r}/\partial v$ están calculadas en $(U(t), V(t))$. Para calcular $\boldsymbol{\rho}'(t)$ derivamos cada componente de $\boldsymbol{\rho}(t)$ mediante la regla de la cadena (teorema 8.8) para obtener

$$\boldsymbol{\rho}'(t) = \nabla X \cdot \boldsymbol{\alpha}'(t)\boldsymbol{i} + \nabla Y \cdot \boldsymbol{\alpha}'(t)\boldsymbol{j} + \nabla Z \cdot \boldsymbol{\alpha}'(t)\boldsymbol{k}, \tag{12.6}$$

donde los vectores gradientes ∇X, ∇Y, y ∇Z están calculados en $(U(t), V(t))$. La ecuación (12.6) puede escribirse en la forma

$$\boldsymbol{\rho}'(t) = \frac{\partial \boldsymbol{r}}{\partial u} U'(t) + \frac{\partial \boldsymbol{r}}{\partial v} V'(t),$$

estando calculadas las derivadas $\partial\boldsymbol{r}/\partial u$ y $\partial\boldsymbol{r}/\partial v$ en $(U(t), V(t))$. Ya que $\partial\boldsymbol{r}/\partial u$ y $\partial\boldsymbol{r}/\partial v$ son perpendiculares en cada punto al producto vectorial $\partial\boldsymbol{r}/\partial u \times \partial\boldsymbol{r}/\partial v$, lo mismo ocurre con $\boldsymbol{\rho}'(t)$. Esto demuestra que $\partial\boldsymbol{r}/\partial u \times \partial\boldsymbol{r}/\partial v$ es normal a C, como queríamos probar. Por esta razón, el vector $\partial\boldsymbol{r}/\partial u \times \partial\boldsymbol{r}/\partial v$ se denomina *normal* a la superficie $\boldsymbol{r}(T)$. En cada punto regular P de $\boldsymbol{r}(T)$ el vector $\partial\boldsymbol{r}/\partial u \times \partial\boldsymbol{r}/\partial v$ es distinto de cero; el plano que pasa por P y tiene este vector como normal se llama *plano tangente* a la superficie en P.

12.4 Ejercicios

En los ejercicios del 1 al 6, eliminar los parámetros u y v para obtener la ecuación cartesiana, probando así que la ecuación vectorial dada representa la superficie que se cita. Calcular también el producto vectorial fundamental $\partial\boldsymbol{r}/\partial u \times \partial\boldsymbol{r}/\partial v$ en función de u y v.

1. *Plano:*
$\boldsymbol{r}(u, v) = (x_0 + a_1u + b_1v)\boldsymbol{i} + (y_0 + a_2u + b_2v)\boldsymbol{j} + (z_0 + a_3u + b_3v)\boldsymbol{k}$.
2. *Paraboloide elíptico:*
$\boldsymbol{r}(u, v) = au \cos v\, \boldsymbol{i} + bu \operatorname{sen} v\, \boldsymbol{j} + u^2\boldsymbol{k}$.
3. *Elipsoide:*
$\boldsymbol{r}(u, v) = a \operatorname{sen} u \cos v\, \boldsymbol{i} + b \operatorname{sen} u \operatorname{sen} v\, \boldsymbol{j} + c \cos u\, \boldsymbol{k}$.
4. *Superficie de revolución:*
$\boldsymbol{r}(u, v) = u \cos v\, \boldsymbol{i} + u \operatorname{sen} v\, \boldsymbol{j} + f(u)\boldsymbol{k}$.
5. *Cilindro:*
$\boldsymbol{r}(u, v) = u\boldsymbol{i} + a \operatorname{sen} v\, \boldsymbol{j} + a \cos v\, \boldsymbol{k}$.
6. *Toro:*
$\boldsymbol{r}(u, v) = (a + b \cos u) \operatorname{sen} v\, \boldsymbol{i} + (a + b \cos u) \cos v\, \boldsymbol{j} + b \operatorname{sen} u\, \boldsymbol{k}$, donde $0 < b < a$.

¿Cuál es el significado geométrico de a y b?

En los ejercicios 7 al 10 calcular la magnitud del producto vectorial fundamental en función de u y v.

7. $\boldsymbol{r}(u, v) = a \operatorname{sen} u \cosh v\, \boldsymbol{i} + b \cos u \cosh v\, \boldsymbol{j} + c \operatorname{senh} v\, \boldsymbol{k}$.
8. $\boldsymbol{r}(u, v) = (u + v)\boldsymbol{i} + (u - v)\boldsymbol{j} + 4v^2\boldsymbol{k}$.
9. $\boldsymbol{r}(u, v) = (u + v)\boldsymbol{i} + (u^2 + v^2)\boldsymbol{j} + (u^3 + v^3)\boldsymbol{k}$.
10. $\boldsymbol{r}(u, v) = u \cos v\, \boldsymbol{i} + u \operatorname{sen} v\, \boldsymbol{j} + \frac{1}{2}u^2 \operatorname{sen} 2v\, \boldsymbol{k}$.

12.5 Área de una superficie paramétrica

Sea $S = \boldsymbol{r}(T)$ una superficie paramétrica representada por la función $\boldsymbol{r}$ definida en una región T del plano uv. En la sección 12.2 vimos que la longitud del producto vectorial fundamental $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ puede interpretarse como un factor de proporcionalidad de áreas. (Véase figura 12.6) Un rectángulo en T de área $\Delta u\, \Delta v$ es aplicado por $\boldsymbol{r}$ sobre un paralelogramo curvilíneo en S con área aproximadamente igual a

$$\left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| \Delta u\, \Delta v .$$

Esta observación sugiere la siguiente definición.

DEFINICIÓN DE ÁREA DE UNA SUPERFICIE PARAMÉTRICA. *El área de S, que representamos por $a(S)$, se define por la integral doble*

$$a(S) = \iint_T \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| du\, dv . \tag{12.7}$$

Dicho de otro modo, para determinar el área de S calculamos primero el producto vectorial fundamental $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ e integramos entonces su longitud en la región T. Cuando $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ está expresado en función de sus componentes, por medio de la ecuación (12.4), tenemos,

$$a(S) = \iint_T \sqrt{\left(\frac{\partial(Y, Z)}{\partial(u, v)}\right)^2 + \left(\frac{\partial(Z, X)}{\partial(u, v)}\right)^2 + \left(\frac{\partial(X, Y)}{\partial(u, v)}\right)^2}\, du\, dv . \tag{12.8}$$

Escrita en esta forma, la integral para el área de una superficie se parece a la inte-

gral usada para calcular la longitud de un arco de curva.(*)

Si S viene dada explícitamente por una ecuación de la forma $z = f(x, y)$ podemos usar x e y como parámetros. El producto vectorial fundamental es el de la ecuación (12.5), de este modo tenemos:

$$\left\| \frac{\partial \boldsymbol{r}}{\partial x} \times \frac{\partial \boldsymbol{r}}{\partial y} \right\| = \left\| -\frac{\partial f}{\partial x}\boldsymbol{i} - \frac{\partial f}{\partial y}\boldsymbol{j} + \boldsymbol{k} \right\| = \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}.$$

En este caso, la integral para calcular el área de la superficie toma la forma:

$$a(S) = \iint_T \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dx\, dy, \tag{12.9}$$

donde la región T es ahora la proyección de S en el plano xy, como se ve en la figura 12.7.

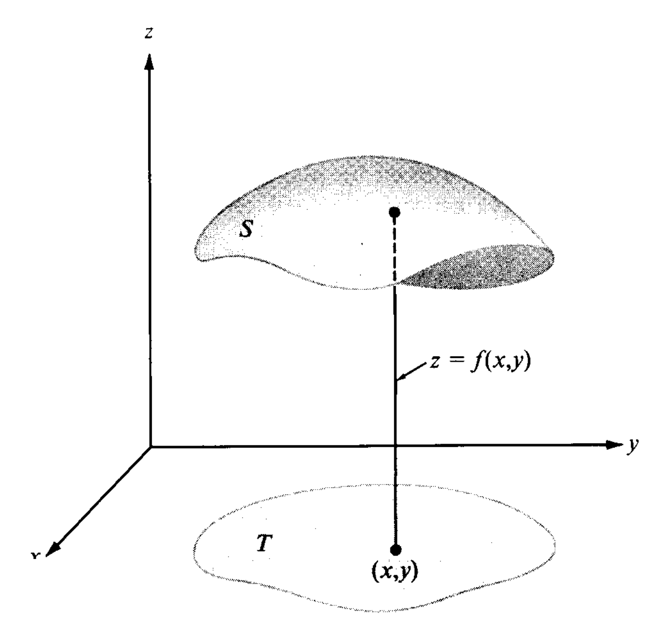


FIGURA 12.7 *Superficie S con una representación explícita, $z = f(x, y)$. La región T es la proyección de S sobre el plano xy.*

(*) Puesto que en la integral (12.7) interviene $\boldsymbol{r}$, el área de una superficie dependerá de la función utilizada para representar la superficie. Cuando comentemos las integrales de superficie demostraremos (en la sección 12.8) que bajo ciertas condiciones generales el área es independiente de la representación paramétrica. El resultado es análogo al del teorema 10.1, en el que se discutió la invariancia de las integrales de línea frente a un cambio de parámetro.

Cuando S está en un plano paralelo al plano xy, la función f es constante, de modo que $\partial f/\partial x = \partial f/\partial y = 0$, y la ecuación (12.9) se convierte en:

$$a(S) = \iint_T dx\, dy\,.$$

Que coincide con la fórmula corriente para el cálculo de áreas de regiones planas.

La ecuación (12.9) puede escribirse en otra forma que nos hace comprender mejor su significado geométrico. En cada punto de S, sea γ el ángulo formado por el vector normal $\boldsymbol{N} = \partial \boldsymbol{r}/\partial x \times \partial \boldsymbol{r}/\partial y$ y el vector unitario coordenado $\boldsymbol{k}$. (Véase figura 12.8.) Puesto que el componente z de $\boldsymbol{N}$ es 1, tenemos:

$$\cos \gamma = \frac{\boldsymbol{N} \cdot \boldsymbol{k}}{\|\boldsymbol{N}\|\, \|\boldsymbol{k}\|} = \frac{1}{\|\boldsymbol{N}\|} = \frac{1}{\left\| \dfrac{\partial \boldsymbol{r}}{\partial x} \times \dfrac{\partial \boldsymbol{r}}{\partial y} \right\|}\,,$$

y por tanto, $\| \partial \boldsymbol{r}/\partial x \times \partial \boldsymbol{r}/\partial y \| = 1/\cos \gamma$. Por consiguiente, la ecuación (12.9) toma la forma:

$$a(S) = \iint_T \frac{1}{\cos \gamma}\, dx\, dy\,. \tag{12.10}$$

Supongamos ahora que S está en un plano no perpendicular al plano xy. En tal caso γ es constante y la ecuación (12.10) establece que área de S = (área de T)/cos γ, o que

(12.11)

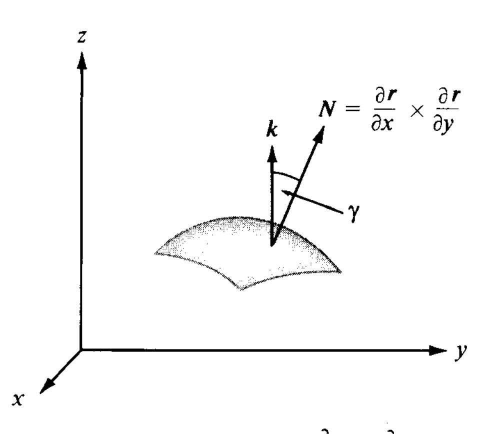


FIGURA 12.8 *La longitud de* $\dfrac{\partial \boldsymbol{r}}{\partial x} \times \dfrac{\partial \boldsymbol{r}}{\partial y}$ *es* $1/\cos \gamma$.

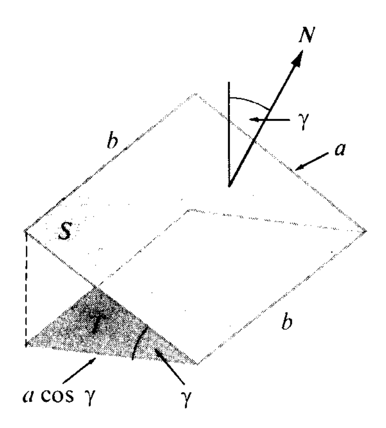


FIGURA 12.9 *La regla del coseno para un rectángulo.*

La ecuación (12.11) se denomina algunas veces *regla del coseno.* Nos dice que si una región S de un plano se proyecta en una región T de otro plano, que forma un ángulo γ con el primero, el área de T es igual al producto de $\cos \gamma$ por el área de S. Esta fórmula es evidentemente cierta cuando S es el rectángulo dibujado en la figura 12.9, pues las distancias en una dirección quedan reducidas por el factor $\cos \gamma$ mientras que las de dirección perpendicular no se alteran en la proyección. La ecuación (12.11) extiende esta propiedad a cualquier región plana S.

Supongamos ahora que S viene dada implícitamente por la ecuación $F(x, y, z) = 0$. Si S puede proyectarse en forma uno a uno sobre el plano xy, la ecuación $F(x, y, z) = 0$ define z como función de x e y, sea ésta $z = f(x, y)$, y las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ están relacionadas con las de F mediante las ecuaciones

$$\frac{\partial f}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z} \quad \text{y} \quad \frac{\partial f}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z}$$

en los puntos en los que $\partial F/\partial z \neq 0$. Sustituyendo esos cocientes en (12.9), encontramos:

$$a(S) = \iint_T \frac{\sqrt{(\partial F/\partial x)^2 + (\partial F/\partial y)^2 + (\partial F/\partial z)^2}}{|\partial F/\partial z|}\, dx\, dy. \tag{12.12}$$

EJEMPLO 1. *Área de un hemisferio.* Consideremos un hemisferio S de radio a y centro en el origen. Disponemos de la representación implícita $x^2 + y^2 + z^2 = a^2$, $z \geqslant 0$; la explícita $z = \sqrt{a^2 - x^2 - y^2}$; y la paramétrica

$$\boldsymbol{r}(u, v) = a \cos u \cos v\, \boldsymbol{i} + a \operatorname{sen} u \cos v\, \boldsymbol{j} + a \operatorname{sen} v\, \boldsymbol{k}. \tag{12.13}$$

Para calcular el área de S a partir de la representación implícita utilizaremos la fórmula (12.12) tomando

$$F(x, y, z) = x^2 + y^2 + z^2 - a^2.$$

Las derivadas parciales de F son $\partial F/\partial x = 2x$, $\partial F/\partial y = 2y$, $\partial F/\partial z = 2z$. El hemisferio S se proyecta en forma uno a uno sobre el disco $D = \{(x, y) \mid x^2 + y^2 \leq a^2\}$ en el plano xy. No podemos aplicar la fórmula (12.12) directamente porque la derivada $\partial F/\partial z$ es nula en la frontera de D. No obstante, dicha derivada no es nula en todo punto interior de D, de modo que podemos considerar un disco menor concéntrico $D(R)$ de radio R, siendo $R < a$. Si $S(R)$ representa la porción correspondiente del hemisferio superior, la (12.12) es ahora aplicable y resulta:

$$\text{area de } S(R) = \iint_{D(R)} \frac{\sqrt{(2x)^2 + (2y)^2 + (2z)^2}}{|2z|}\, dx\, dy$$

$$= \iint_{D(R)} \frac{a}{z}\, dx\, dy = a \iint_{D(R)} \frac{1}{\sqrt{a^2 - x^2 - y^2}}\, dx\, dy .$$

La última integral puede calcularse fácilmente en coordenadas polares, obteniendo:

$$\text{area de } S(R) = a \int_0^{2\pi} \left[\int_0^R \frac{1}{\sqrt{a^2 - r^2}}\, r\, dr \right] d\theta = 2\pi a(a - \sqrt{a^2 - R^2}).$$

Cuando $R \to a$ resulta como límite $2\pi a^2$.

Puede eludirse el paso al límite en el cálculo precedente empleando la representación paramétrica (12.13). Los cálculos del ejemplo 2 de la sección 12.2 muestran que

$$\left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| = \|a \cos v\, \boldsymbol{r}(u, v)\| = a^2\, |\cos v| .$$

Por lo tanto, podemos aplicar la fórmula (12.7), tomando como región T el rectángulo $[0, 2\pi] \times [0, \frac{1}{2}\pi]$. Obtenemos:

$$a(S) = a^2 \iint_T |\cos v|\, du\, dv = a^2 \int_0^{2\pi} \left[\int_0^{\pi/2} \cos v\, dv \right] du = 2\pi a^2 .$$

EJEMPLO 2. *Otro teorema de Pappus.* Uno de los teoremas de Pappus establece que una superficie de revolución, obtenida por la rotación de una curva plana de longitud L alrededor de un eje situado en el plano de la curva, tiene por área $2\pi Lh$ siendo h la distancia desde el centro de gravedad de la curva al eje de rotación. Usaremos la fórmula (12.7) para demostrar este teorema en un caso especial de superficie de revolución.

Supongamos una curva C, inicialmente en el plano xz, que gira alrededor del eje z. Sea $z = f(x)$, su ecuación en el plano xz, donde $a \leqslant x \leqslant b$, $a \geqslant 0$. La superficie de revolución S así engendrada puede representarse por la ecuación vectorial

$$\boldsymbol{r}(u, v) = u \cos v\, \boldsymbol{i} + u \operatorname{sen} v\, \boldsymbol{j} + f(u) \boldsymbol{k} ,$$

donde $(u, v) \in [a, b] \times [0, 2\pi]$. Los parámetros u y v pueden interpretarse como el radio y ángulo polares, como se ve en la figura 12.10. Si $a \leqslant u \leqslant b$, todos los puntos (x, y, z) situados a una misma distancia u del eje z, tienen la misma coordenada z, $f(u)$, de manera que están todos en la superficie. El producto vectorial fundamental de esta representación es:

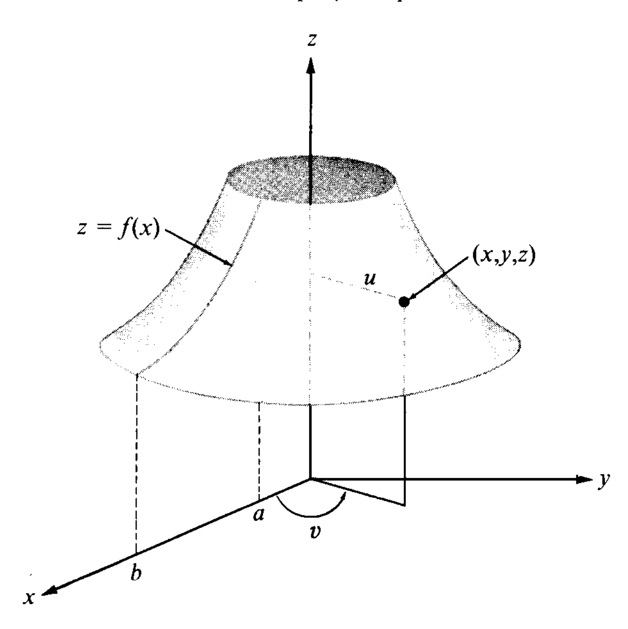


FIGURA 12.10 *Área de una superficie de revolución determinada mediante el teorema de Pappus.*

$$\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \cos v & \operatorname{sen} v & f'(u) \\ -u \operatorname{sen} v & u \cos v & 0 \end{vmatrix} = -uf'(u)\cos v\, \boldsymbol{i} - uf'(u) \operatorname{sen} v\, \boldsymbol{j} + u\boldsymbol{k},$$

y por tanto,

$$\left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| = u\sqrt{1 + [f'(u)]^2}\,.$$

Por consiguiente, la fórmula (12.7) se convierte en:

$$a(S) = \int_0^{2\pi} \left[\int_a^b u\sqrt{1 + [f'(u)]^2}\, du \right] dv = 2\pi \int_a^b u\sqrt{1 + [f'(u)]^2}\, du\,.$$

La integral que resulta puede expresarse como $\int_C x\, ds$, que es una integral de línea respecto a la longitud del arco tomada a lo largo de C. Como tal, vale $\bar{x}L$, donde $\bar{x}$ es la coordenada x del centro de gravedad de C y L la longitud de C. (Véase sección 10.8.) Luego el área de S es $2\pi L\bar{x}$. Esto demuestra el teorema de Pappus en la superficie del ejemplo considerado.

12.6 Ejercicios

1. Sea S un paralelogramo de lados no paralelos a ningún eje coordenado. Sean S_1, S_2 y S_3 las áreas de las proyecciones de S sobre los planos coordenados. Demostrar que el área de S es $\sqrt{S_1^2 + S_2^2 + S_3^2}$.
2. Calcular el área de la región que en el plano $x + y + z = a$ determina el cilindro $x^2 + y^2 = a^2$.
3. Calcular el área de la porción de esfera $x^2 + y^2 + z^2 = a^2$ interior al cilindro $x^2 + y^2 = ay$, siendo $a > 0$.
4. Calcular el área de la porción de superficie $z^2 = 2xy$ que se proyecta en el primer cuadrante del plano xy y limitada por los planos $x = 2$ e $y = 1$.
5. Dada una superficie S de ecuación vectorial

$$\boldsymbol{r}(u, v) = u \cos v\, \boldsymbol{i} + u \operatorname{sen} v\, \boldsymbol{j} + u^2 \boldsymbol{k},$$

donde $0 \le u \le 4$ y $0 \le v \le 2\pi$.
a) Demostrar que S es una porción de cuádrica. Identificar dicha cuádrica, dibujarla e indicar el significado geométrico de los parámetros u y v en la superficie.
b) Calcular el producto vectorial fundamental $\partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ en función de u y v.
c) El área de S es $\pi(65\sqrt{65} - 1)/n$, donde n es entero. Calcular el valor de n.
6. Calcular el área de la porción de superficie cónica $x^2 + y^2 = z^2$ situada por encima del plano xy y limitada por la esfera $x^2 + y^2 + z^2 = 2ax$.
7. Calcular el área de la porción de superficie cónica $x^2 + y^2 = z^2$ situada entre los dos planos $z = 0$ y $x + 2z = 3$.
8. Calcular el área de la porción de paraboloide $x^2 + y^2 = 2ay$ cortada por el plano $y = a$.
9. Calcular el área del toro de ecuación vectorial

$$\boldsymbol{r}(u, v) = (a + b \cos u) \operatorname{sen} v\, \boldsymbol{i} + (a + b \cos u) \cos v\, \boldsymbol{j} + b \operatorname{sen} u\, \boldsymbol{k},$$

donde $0 < b < a$ y $0 \le u \le 2\pi$, $0 \le v \le 2\pi$. Utilizar el teorema de Pappus para comprobar el resultado.
10. Una esfera está inscrita en un cilindro circular recto. La esfera es cortada por dos planos paralelos perpendiculares al eje del cilindro. Demostrar que las porciones de la esfera y del cilindro comprendidas entre esos dos planos tienen la misma área.
11. Sea T el disco unidad en el plano uv, $T = \{(u, v) \mid u^2 + v^2 \le 1\}$, y pongamos

$$\boldsymbol{r}(u, v) = \frac{2u}{u^2 + v^2 + 1}\boldsymbol{i} + \frac{2v}{u^2 + v^2 + 1}\boldsymbol{j} + \frac{u^2 + v^2 - 1}{u^2 + v^2 + 1}\boldsymbol{k}.$$

a) Determinar la imagen por $\boldsymbol{r}$ de cada uno de los siguientes conjuntos: la circunferencia unidad $u^2 + v^2 = 1$; el intervalo $-1 \le u \le 1$; la parte de la recta $u = v$ situada en T.
b) La superficie $S = \boldsymbol{r}(T)$ es muy conocida. Decir cuál es y dibujarla.
c) Determinar la imagen por $\boldsymbol{r}$ del plano uv. Indicar con un gráfico en el espacio xyz los significados geométricos de los parámetros u y v.

12.7 Integrales de superficie

En muchos aspectos, las integrales de superficie son análogas a las integrales de línea. Definimos las integrales de línea mediante una representación paramétrica de la curva. Análogamente, definiremos las integrales de superficie en función de una representación paramétrica de la superficie. Demostraremos luego que en ciertas condiciones generales el valor de la integral es independiente de la representación.

DEFINICIÓN DE INTEGRAL DE SUPERFICIE. *Sea $S = \boldsymbol{r}(T)$ una superficie paramétrica descrita por una función diferenciable $\boldsymbol{r}$ definida en una región T del plano uv, y sea f un campo escalar definido y acotado en S. La integral de superficie de f sobre S se representa con el símbolo $\iint_{\boldsymbol{r}(T)} f\,dS$ [o por $\iint_S f(x, y, z)\,dS$], y está definida por la ecuación*

$$\iint_{\boldsymbol{r}(T)} f\,dS = \iint_T f[\boldsymbol{r}(u, v)] \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| du\,dv \tag{12.14}$$

siempre que exista la integral doble del segundo miembro.

Los ejemplos que siguen ilustran algunas aplicaciones de las integrales de superficie.

EJEMPLO 1. *Área de una superficie.* Cuando $f = 1$, la ecuación (12.14) se transforma en

$$\iint_{\boldsymbol{r}(T)} dS = \iint_T \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| du\,dv\,.$$

La integral doble del segundo miembro es la que se usó en la sección 12.5 para definir el área de una superficie. Así pues, el área de S es igual a la integral de superficie $\iint_{\boldsymbol{r}(T)} dS$. Por este motivo, el símbolo dS se llama algunas veces «elemento de área de la superficie», y la integral de superficie $\iint_{\boldsymbol{r}(T)} f\,dS$ se lee integral de f respecto al elemento de área, extendida a la superficie $\boldsymbol{r}(T)$.

EJEMPLO 2. *Centro de gravedad. Momento de inercia.* Si el campo escalar f se interpreta como la densidad (masa por unidad de área) de una lámina delgada

adaptada a la superficie S, la masa total m de la superficie se define por la fórmula

$$m = \iint_S f(x, y, z)\, dS .$$

Su centro de gravedad es el punto $(\bar{x}, \bar{y}, \bar{z})$ determinado por las fórmulas

$$\bar{x}m = \iint_S xf(x, y, z)\, dS , \qquad \bar{y}m = \iint_S yf(x, y, z)\, dS , \qquad \bar{z}m = \iint_S zf(x, y, z)\, dS .$$

El momento de inercia I_L de S alrededor de un eje L viene dado por

$$I_L = \iint_S \delta^2(x, y, z) f(x, y, z)\, dS ,$$

donde $\delta(x, y, z)$ representa la distancia de un punto genérico (x, y, z) de S a la recta L.

Como ejemplo, determinemos el centro de gravedad de la superficie de una semiesfera uniforme de radio a. Utilicemos la representación paramétrica

$$\boldsymbol{r}(u, v) = a \cos u \cos v\, \boldsymbol{i} + a \operatorname{sen} u \cos v\, \boldsymbol{j} + a \operatorname{sen} v\, \boldsymbol{k} ,$$

en la que $(u, v) \in [0, 2\pi] \times [0, \frac{1}{2}\pi]$. Esta representación ya fue discutida en el ejemplo 2 de la sección 12.2, donde encontramos que la magnitud del producto vectorial fundamental es $a^2 |\cos v|$. En este ejemplo la densidad f es constante, pongamos $f = c$, y la masa m es $2\pi a^2 c$, c veces el área de S. Debido a la simetría, las coordenadas $\bar{x}$ e $\bar{y}$ del centro de gravedad son 0. La coordenada $\bar{z}$ viene dada por

$$\begin{aligned} \bar{z}m &= c \iint_S z\, dS = c \iint_T a \operatorname{sen} v \cdot a^2 |\cos v|\, du\, dv \\ &= 2\pi a^3 c \int_0^{\pi/2} \operatorname{sen} v \cos v\, dv = \pi a^3 c = \frac{a}{2} m , \end{aligned}$$

de modo que $\bar{z} = a/2$.

EJEMPLO 3. *Flujo de fluido a través de una superficie.* Imaginemos que un fluido es una colección de puntos llamados *partículas*. A cada partícula (x, y, z) asignamos un vector $V(x, y, z)$ que representa su velocidad. Este es el campo de velocidad de la corriente. El campo de velocidad puede o no cambiar con el tiempo. Consideraremos tan sólo corrientes *estacionarias*, esto es, corrientes para las

que la velocidad $V(x, y, z)$ depende únicamente de la posición de la partícula y no del tiempo.

Designemos con $\rho(x, y, z)$ la densidad (masa por unidad de volumen) del fluido en el punto (x, y, z). Si el fluido es incompresible la dentidad ρ será constante en todo el fluido. Para un fluido compresible, tal como un gas, la densidad puede variar de un punto a otro. En cualquier caso, la densidad es un campo escalar asociado a la corriente. El producto de la densidad por la velocidad la representamos por $\boldsymbol{F}$; esto es,

$$\boldsymbol{F}(x, y, z) = \rho(x, y, z)\boldsymbol{V}(x, y, z).$$

Este es un campo vectorial llamado *densidad de flujo* de la corriente. El vector $\boldsymbol{F}(x, y, z)$ tiene la misma dirección que la velocidad, y su longitud tiene las dimensiones

$$\frac{\text{masa}}{\text{unidad de volumen}} \cdot \frac{\text{distancia}}{\text{unidad de tiempo}} = \frac{\text{masa}}{\text{(unidad de área)(unidad de tiempo)}}$$

Dicho de otro modo, el vector densidad de flujo $\boldsymbol{F}(x, y, z)$ nos dice cuánta masa de fluido circula por el punto (x, y, z) en la dirección de $\boldsymbol{V}(x, y, z)$, por unidad de área y de tiempo.

Sea $S = \boldsymbol{r}(T)$ una superficie paramétrica simple. En cada punto regular de S designemos con $\boldsymbol{n}$ el vector unitario normal que tenga el mismo sentido que el producto vectorial fundamental. Esto es,

$$\text{(12.15)} \qquad \boldsymbol{n} = \frac{\dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v}}{\left\| \dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v} \right\|}.$$

El producto escalar $\boldsymbol{F} \cdot \boldsymbol{n}$ representa el componente del vector densidad de flujo en la dirección de $\boldsymbol{n}$. La masa de fluido que pasa a través de S en la unidad de tiempo en la dirección de $\boldsymbol{n}$ se define con la integral de superficie

$$\iint_{\boldsymbol{r}(T)} \boldsymbol{F} \cdot \boldsymbol{n} \, dS = \iint_T \boldsymbol{F} \cdot \boldsymbol{n} \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| du \, dv .$$

12.8 Cambio de representación paramétrica

Volvamos nuevamente a la discusión de la independencia de las integrales de superficie de los cambios de representación paramétrica. Supongamos que una

función $\boldsymbol{r}$ aplica una región A del plano uv sobre una superficie paramétrica $\boldsymbol{r}(A)$. Supongamos también que A es la imagen de una región B del plano st a través de una aplicación uno a uno derivable con continuidad $\boldsymbol{G}$ dada por

$$\boldsymbol{G}(s, t) = U(s, t)\boldsymbol{i} + V(s, t)\boldsymbol{j} \quad \text{si} \quad (s, t) \in B. \tag{12.16}$$

Consideremos la función $\boldsymbol{R}$ definida en B mediante la ecuación

$$\boldsymbol{R}(s, t) = \boldsymbol{r}[\boldsymbol{G}(s, t)]. \tag{12.17}$$

(Véase figura 12.11.) Dos funciones $\boldsymbol{r}$ y $\boldsymbol{R}$ así relacionadas se denominarán *regularmente equivalentes.* Funciones regularmente equivalentes representan la misma superficie. Esto es, $\boldsymbol{r}(A)$ y $\boldsymbol{R}(B)$ como conjuntos de puntos son idénticos. (Esto resulta inmediatamente del hecho de ser $\boldsymbol{G}$ uno a uno.) El siguiente teorema trata de la relación entre sus productos vectoriales fundamentales.

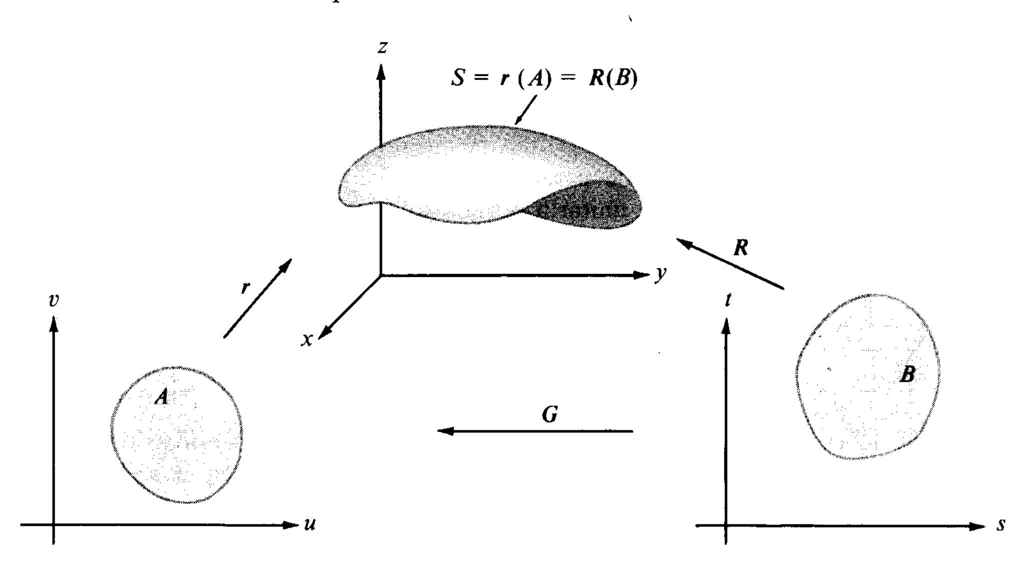


FIGURA 12.11 *Dos representaciones paramétricas de la misma superficie.*

TEOREMA 12.1. *Sean $\boldsymbol{r}$ y $\boldsymbol{R}$ dos funciones regularmente equivalentes ligadas por la ecuación* (12.17), *donde $\boldsymbol{G} = U\boldsymbol{i} + V\boldsymbol{j}$ es una aplicación, uno a uno y derivable con continuidad, de una región B del plano st sobre una región A del plano uv dada por la ecuación* (12.16). *Tenemos entonces*

$$\frac{\partial \boldsymbol{R}}{\partial s} \times \frac{\partial \boldsymbol{R}}{\partial t} = \left(\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}\right)\frac{\partial(U, V)}{\partial(s, t)}, \tag{12.18}$$

donde las derivadas parciales $\partial \boldsymbol{r}/\partial u$ *y* $\partial \boldsymbol{r}/\partial v$ *están calculadas en el punto* $(U(s,t), V(s,t))$. *Dicho de otro modo, el producto vectorial fundamental de* $\boldsymbol{R}$ *es igual al de* $\boldsymbol{r}$, *multiplicado por el determinante jacobiano de la aplicación* $\boldsymbol{G}$.

Demostración. Las derivadas $\partial \boldsymbol{R}/\partial s$ y $\partial \boldsymbol{R}/\partial t$ pueden calcularse por derivación de (12.17). Si aplicamos la regla de la cadena (teorema 8.8) a cada componente de $\boldsymbol{R}$ y reordenamos los términos, encontramos que

$$\frac{\partial \boldsymbol{R}}{\partial s} = \frac{\partial \boldsymbol{r}}{\partial u}\frac{\partial U}{\partial s} + \frac{\partial \boldsymbol{r}}{\partial v}\frac{\partial V}{\partial s} \qquad \text{y} \qquad \frac{\partial \boldsymbol{R}}{\partial t} = \frac{\partial \boldsymbol{r}}{\partial u}\frac{\partial U}{\partial t} + \frac{\partial \boldsymbol{r}}{\partial v}\frac{\partial V}{\partial t},$$

donde las derivadas $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ están calculadas en $(U(s,t), V(s,t))$. Multiplicando vectorialmente esas dos ecuaciones obtenemos:

$$\frac{\partial \boldsymbol{R}}{\partial s} \times \frac{\partial \boldsymbol{R}}{\partial t} = \left(\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}\right)\left(\frac{\partial U}{\partial s}\frac{\partial V}{\partial t} - \frac{\partial U}{\partial t}\frac{\partial V}{\partial s}\right) = \left(\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}\right)\frac{\partial(U,V)}{\partial(s,t)}.$$

Como queríamos demostrar.

La invariancia de las integrales de superficie frente a las representaciones paramétricas regularmente equivalentes resulta ahora como consecuencia inmediata del teorema 12.1.

TEOREMA 12.2. *Si* $\boldsymbol{r}$ *y* $\boldsymbol{R}$ *son dos funciones regularmente equivalentes, como las descritas en el teorema* 12.1 *y si la integral de superficie* $\iint_{\boldsymbol{r}(A)} f\, dS$ *existe, entonces también existe* $\iint_{\boldsymbol{R}(B)} f\, dS$ *y tenemos*

$$\iint_{\boldsymbol{r}(A)} f\, dS = \iint_{\boldsymbol{R}(B)} f\, dS.$$

Demostración. Según la definición de integral de superficie tenemos

$$\iint_{\boldsymbol{r}(A)} f\, dS = \iint_{A} f[\boldsymbol{r}(u,v)] \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| du\, dv.$$

Utilicemos ahora la aplicación $\boldsymbol{G}$ del teorema 12.1 para transformar esa integral en una integral doble extendida a la región B del plano st. La fórmula de transformación de las integrales dobles establece que

$$\iint_{A} f[\boldsymbol{r}(u,v)] \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| du\, dv = \iint_{B} f\{\boldsymbol{r}[\boldsymbol{G}(s,t)]\} \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| \left| \frac{\partial(U,V)}{\partial(s,t)} \right| ds\, dt,$$

en donde las derivadas $\partial \boldsymbol{r}/\partial u$ y $\partial \boldsymbol{r}/\partial v$ del segundo miembro están calculadas en $(U(s, t), V(s, t))$. En virtud de la ecuación (12.18), la integral sobre B es igual al

$$\iint_B f[\boldsymbol{R}(s, t)] \left\| \frac{\partial \boldsymbol{R}}{\partial s} \times \frac{\partial \boldsymbol{R}}{\partial t} \right\| ds\, dt .$$

Esto es la definición de la integral de superficie $\iint_{R(B)} f\, dS$. Con lo cual queda completa la demostración.

12.9 Otras notaciones para las integrales de superficie

Si $S = \boldsymbol{r}(T)$ es una superficie paramétrica, el producto vectorial fundamental $\boldsymbol{N} = \partial \boldsymbol{r}/\partial u \times \partial \boldsymbol{r}/\partial v$ es normal a S en cada punto regular de la superficie. En cada uno de tales puntos existen *dos* vectores normales unitarios, uno $\boldsymbol{n}_1$ que tiene la misma dirección que $\boldsymbol{N}$, y otro $\boldsymbol{n}_2$ que tiene dirección apuesta. Así pues,

$$\boldsymbol{n}_1 = \frac{\boldsymbol{N}}{\|\boldsymbol{N}\|} \qquad \text{y} \qquad \boldsymbol{n}_2 = -\boldsymbol{n}_1 .$$

Sea $\boldsymbol{n}$ una de las dos normales $\boldsymbol{n}_1$ ó $\boldsymbol{n}_2$. Sea $\boldsymbol{F}$ un campo vectorial definido en S y supongamos que la integral de superficie $\iint_S \boldsymbol{F} \cdot \boldsymbol{n}\, dS$ exista. Podemos entonces escribir

$$(12.19) \qquad \iint_S \boldsymbol{F} \cdot \boldsymbol{n}\, dS = \iint_T \boldsymbol{F}[\boldsymbol{r}(u, v)] \cdot \boldsymbol{n}(u, v) \left\| \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} \right\| du\, dv$$

$$= \pm \iint_T \boldsymbol{F}[\boldsymbol{r}(u, v)] \cdot \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}\, du\, dv ,$$

donde el signo $+$ se usa si $\boldsymbol{n} = \boldsymbol{n}_1$ y el signo $-$ si $\boldsymbol{n} = \boldsymbol{n}_2$.

Supongamos ahora que expresamos $\boldsymbol{F}$ y $\boldsymbol{r}$ en función de sus coordenadas,

$$\boldsymbol{F}(x, y, z) = P(x, y, z)\boldsymbol{i} + Q(x, y, z)\boldsymbol{j} + R(x, y, z)\boldsymbol{k}$$

y

$$\boldsymbol{r}(u, v) = X(u, v)\boldsymbol{i} + Y(u, v)\boldsymbol{j} + Z(u, v)\boldsymbol{k} .$$

El producto vectorial fundamental de $\boldsymbol{r}$ viene entonces dado por

$$\boldsymbol{N} = \frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} = \frac{\partial(Y, Z)}{\partial(u, v)}\boldsymbol{i} + \frac{\partial(Z, X)}{\partial(u, v)}\boldsymbol{j} + \frac{\partial(X, Y)}{\partial(u, v)}\boldsymbol{k} .$$

Si $\boldsymbol{n} = \boldsymbol{n}_1$, la ecuación (12.19) se transforma en

$$\iint_S \boldsymbol{F} \cdot \boldsymbol{n}\, dS = \iint_T P[\boldsymbol{r}(u, v)] \frac{\partial(Y, Z)}{\partial(u, v)}\, du\, dv$$

$$+ \iint_T Q[\boldsymbol{r}(u, v)] \frac{\partial(Z, X)}{\partial(u, v)}\, du\, dv + \iint_T R[\boldsymbol{r}(u, v)] \frac{\partial(X, Y)}{\partial(u, v)}\, du\, dv\,; \tag{12.20}$$

si $\boldsymbol{n} = \boldsymbol{n}_2$, cada una de las integrales dobles del segundo miembro debe reemplazarse por su opuesta. Frecuentemente la suma de las integrales dobles del segundo miembro se escribe más brevemente poniendo

$$\iint_S P(x, y, z)\, dy \wedge dz + \iint_S Q(x, y, z)\, dz \wedge dx + \iint_S R(x, y, z)\, dx \wedge dy\,, \tag{12.21}$$

o aun con mayor simplicidad

$$\iint_S P\, dy \wedge dz + Q\, dz \wedge dx + R\, dx \wedge dy\,. \tag{12.22}$$

Las integrales que aparecen en (12.21) y (12.22) pueden expresarse como integrales de superficie. Así, por ejemplo, la integral de superficie $\iint_S P\, dy \wedge dz$ está definida por la ecuación

$$\iint_S P\, dy \wedge dz = \iint_T P[\boldsymbol{r}(u, v)] \frac{\partial(Y, Z)}{\partial(u, v)}\, du\, dv\,. \tag{12.23}$$

Esta notación está sugerida por la fórmula del cambio de variables en una integral doble.

A pesar del parecido en la notación, la integral del primer miembro de (12.23) *no* es una integral doble. Ante todo, P es una función de tres variables. Debemos también tener en cuenta el orden en el que aparecen los símbolos dy y dz en la integral de superficie, debido a que

$$\frac{\partial(Y, Z)}{\partial(u, v)} = -\frac{\partial(Z, Y)}{\partial(u, v)},$$

y por tanto,

$$\iint_S P\, dy \wedge dz = -\iint_S P\, dz \wedge dy\,.$$

Con tal notación, la fórmula (12.20) se puede escribir:

(12.24) $$\iint_S \boldsymbol{F}\cdot\boldsymbol{n}\,dS = \iint_S P\,dy\wedge dz + Q\,dz\wedge dx + R\,dx\wedge dy$$

si $\boldsymbol{n} = \boldsymbol{n}_1$. Si $\boldsymbol{n} = \boldsymbol{n}_2$ la integral del segundo miembro debe reemplazarse por su opuesta. Esta fórmula se parece a la de las integrales de línea:

$$\int_C \boldsymbol{F}\cdot d\boldsymbol{\alpha} = \int_C P\,dx + Q\,dy + R\,dz\,.$$

Si el vector normal unitario $\boldsymbol{n}$ se expresa en función de sus cosenos directores,

$$\boldsymbol{n} = \cos\alpha\,\boldsymbol{i} + \cos\beta\,\boldsymbol{j} + \cos\gamma\,\boldsymbol{k}\,,$$

entonces $\boldsymbol{F}\cdot\boldsymbol{n} = P\cos\alpha + Q\cos\beta + R\cos\gamma$, y podemos escribir

$$\iint_S \boldsymbol{F}\cdot\boldsymbol{n}\,dS = \iint_S (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,dS\,.$$

Esta ecuación es válida tanto si $\boldsymbol{n} = \boldsymbol{n}_1$ como si $\boldsymbol{n} = \boldsymbol{n}_2$. Los cosenos directores dependerán de la elección de la normal. Si $\boldsymbol{n} = \boldsymbol{n}_1$ podemos utilizar (12.24) para escribir

(12.25)

$$\iint_S (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,dS = \iint_S P\,dy\wedge dz + Q\,dz\wedge dx + R\,dx\wedge dy\,.$$

Si, en cambio, $\boldsymbol{n} = \boldsymbol{n}_2$, tenemos

(12.26)

$$\iint_S (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,dS = -\iint_S P\,dy\wedge dz + Q\,dz\wedge dx + R\,dx\wedge dy$$

12.10 Ejercicios

1. Sea S la semiesfera $x^2+y^2+z^2=1$, $z\geq 0$, y $\boldsymbol{F}(x,y,z)=x\boldsymbol{i}+y\boldsymbol{j}$. Sea $\boldsymbol{n}$ el vector normal unitario exterior a S. Calcular el valor de la integral de superficie $\iint_S \boldsymbol{F}\cdot\boldsymbol{n}\,dS$, empleando:
   a) la representación vectorial $\boldsymbol{r}(u,v) = \operatorname{sen} u\cos v\,\boldsymbol{i} + \operatorname{sen} u\operatorname{sen} v\,\boldsymbol{j} + \cos u\,\boldsymbol{k}$,
   b) la representación explícita $z = \sqrt{1-x^2-y^2}$.

2. Demostrar que el momento de inercia de un recipiente esférico alrededor de un diámetro es igual a $\frac{2}{3}ma^2$, siendo m la masa del recipiente y a su radio.
3. Hallar el centro de gravedad de la porción de superficie esférica homogénea $x^2 + y^2 + z^2 = a^2$ situada sobre el primer cuadrante del plano xy.
4. Sea S la porción de plano limitada por un triángulo de vértices $(1, 0, 0), (0, 1, 0)$, y $(0, 0, 1)$, y sea $\boldsymbol{F}(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$. Representamos con $\boldsymbol{n}$ la normal unitaria a S que tiene componente z no negativa. Calcular la integral de superficie $\iint_S \boldsymbol{F} \cdot \boldsymbol{n}\, dS$ utilizando:
   a) la representación vectorial $\boldsymbol{r}(u, v) = (u + v)\boldsymbol{i} + (u - v)\boldsymbol{j} + (1 - 2u)\boldsymbol{k}$,
   b) una representación explícita de la forma $z = f(x, y)$.
5. Sea S una superficie paramétrica dada en la forma explícita $z = f(x, y)$, donde (x, y) varía en una región plana T, proyección de S en el plano xy. Sean $\boldsymbol{F} = P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k}$ y $\boldsymbol{n}$ la normal unitaria a S de componente z no negativa. Emplear la representación paramétrica $\boldsymbol{r}(x, y) = x\boldsymbol{i} + y\boldsymbol{j} + f(x, y)\boldsymbol{k}$ y demostrar que

$$\iint_S \boldsymbol{F} \cdot \boldsymbol{n}\, dS = \iint_T \left(-P \frac{\partial f}{\partial x} - Q \frac{\partial f}{\partial y} + R \right) dx\, dy,$$

   donde P, Q y R están calculadas en $(x, y, f(x, y))$.
6. Sea S la misma superficie del ejercicio 5, y sea φ un campo escalar. Demostrar que

a) $$\iint_S \varphi(x, y, z)\, dS = \iint_T \varphi[x, y, f(x, y)] \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dx\, dy.$$

b) $$\iint_S \varphi(x, y, z)\, dy \wedge dz = -\iint_T \varphi[x, y, f(x, y)] \frac{\partial f}{\partial x}\, dx\, dy.$$

c) $$\iint_S \varphi(x, y, z)\, dz \wedge dx = -\iint_T \varphi[x, y, f(x, y)] \frac{\partial f}{\partial y}\, dx\, dy.$$

7. Si S es la superficie de la esfera $x^2 + y^2 + z^2 = a^2$, calcular el valor de la integral de superficie

$$\iint_S xz\, dy \wedge dz + yz\, dz \wedge dx + x^2\, dx \wedge dy.$$

   Elegir una representación para la que el producto vectorial fundamental tenga la dirección de la normal exterior.
8. El cilindro $x^2 + y^2 = 2x$ recorta una porción de superficie S en la hoja superior del cono $x^2 + y^2 = z^2$. Calcular la integral de superficie

$$\iint_S (x^4 - y^4 + y^2z^2 - z^2x^2 + 1)\, dS.$$

9. Un recipiente esférico homogéneo de radio a está cortado por una hoja de un cono

circular recto cuyo vértice está en el centro de la esfera. Si el ángulo en el vértice del cono es α, siendo $0 < \alpha < \pi$, determinar (en función a a y α) el centro de gravedad de la porción del recipiente esférico que es interior al cono.

10. Una hoja de papel rectangular homogénea de base $2\pi a$ y altura h se arrolla formando una superficie cilíndrica S de radio a. Calcular el momento de inercia de S alrededor de un eje que contiene un diámetro de la base circular.
11. En relación con el ejercicio 10, calcular el momento de inercia de S alrededor de un eje situado en el plano de la base y que es tangente al borde circular de la base.
12. Un flujo de fluido tiene como vector densidad de flujo $\boldsymbol{F}(x, y, z) = x\boldsymbol{i} - (2x + y)\boldsymbol{j} + z\boldsymbol{k}$. Designemos con S el hemisferio $x^2 + y^2 + z^2 = 1$, $z \geq 0$, y con $\boldsymbol{n}$ la normal unitaria orientada hacia el exterior de la esfera. Calcular la masa de fluido que atraviesa S en la unidad de tiempo en la dirección de $\boldsymbol{n}$.
13. Resolver el ejercicio 12 si S contiene también la base plana del hemisferio. En la base inferior la normal unitaria es $-\boldsymbol{k}$.
14. Sea S la porción del plano $x+y+z=t$ determinada en él por la esfera $x^2+y^2+z^2=1$. Sea $\varphi(x, y, z) = 1 - x^2 - y^2 - z^2$ si (x, y, z) es interior a dicha esfera, y $\varphi(x, y, z) = 0$ en cualquier otro caso. Demostrar que

$$\iint_S \varphi(x, y, z)\, dS = \begin{cases} \dfrac{\pi}{18}(3 - t^2)^2 & \text{si } |t| \leq \sqrt{3}, \\ 0 & \text{si } |t| > \sqrt{3}. \end{cases}$$

[*Indicación.* Tomar unos ejes coordenados (x_1, y_1, z_1) con el eje z_1 normal al plano $x + y + z = t$. Usar luego coordenadas polares en el plano x_1y_1 como parámetros para S.]

12.11 Teorema de Stokes

El resto de este capítulo está especialmente dedicado a dos generalizaciones del segundo teorema fundamental del Cálculo a las integrales de superficie. Se conocen, respectivamente, con las denominaciones de *Teorema de Stokes* (*) y *Teorema de la divergencia.* Esta sección trata del primero, y en la sección 12.19 se expondrá el segundo.

El teorema de Stokes es una extensión directa del teorema de Green el cual establece que

$$\iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \int_C P\, dx + Q\, dy,$$

en donde S es una región plana limitada por una curva cerrada C recorrida en sentido positivo (contrario al de las agujas del reloj). El teorema de Stokes relaciona una integral de superficie con una integral de línea y puede enunciarse así:

(*) En honor de G. G. Stokes (1819-1903), matemático irlandés que consiguió contribuciones fundamentales en la Hidrodinámica y en la Óptica.

TEOREMA 12.3. TEOREMA DE STOKES. *Supongamos que S es una superficie paramétrica simple regular, $S = \boldsymbol{r}(T)$, siendo T una región del plano uv limitada por una curva de Jordan regular a trozos Γ. (Ver figura 12.12.) Supongamos también que $\boldsymbol{r}$ es una aplicación uno a uno cuyos componentes tienen derivadas parciales segundas continuas en un cierto conjunto abierto que contenga $T \cup \Gamma$. Designemos con C la imagen de Γ por $\boldsymbol{r}$, y sean P, Q y R campos escalares derivables con continuidad en S. Tenemos entonces*

$$\iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy \wedge dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dz \wedge dx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy$$

$$(12.27) \qquad = \int_C P\,dx + Q\,dy + R\,dz\,.$$

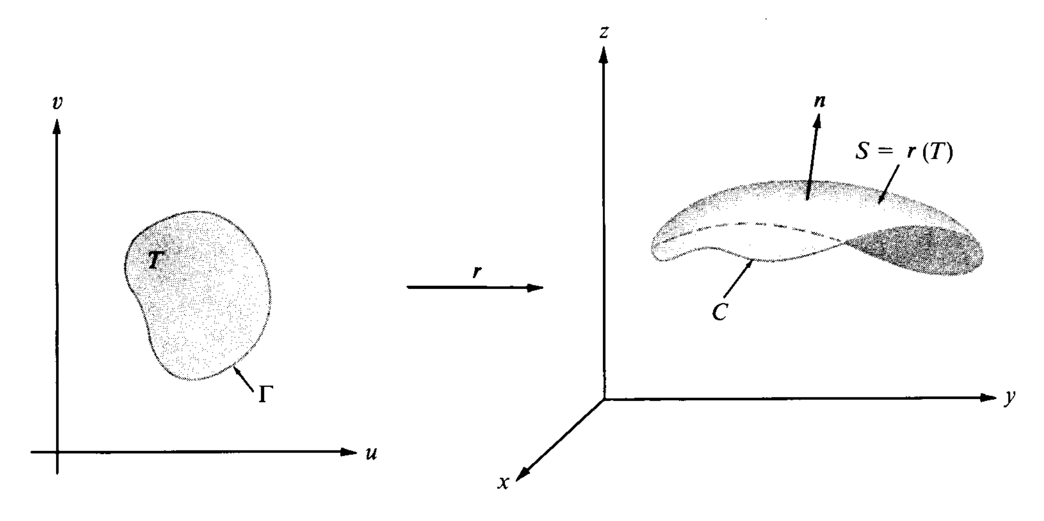


FIGURA 12.12 *Ejemplo de superficie a la que es aplicable el teorema de Stokes.*

La curva Γ se recorre en sentido positivo (contrario al de las agujas del reloj) y la curva C en el sentido que resulte de aplicar a Γ la función $\boldsymbol{r}$.

Demostración. Para demostrar el teorema basta establecer las tres fórmulas siguientes,

$$(12.28) \qquad \int_C P\,dx = \iint_S \left(-\frac{\partial P}{\partial y}\,dx \wedge dy + \frac{\partial P}{\partial z}\,dz \wedge dx\right),$$

$$\int_C Q\,dy = \iint_S \left(-\frac{\partial Q}{\partial z}\,dy \wedge dz + \frac{\partial Q}{\partial x}\,dx \wedge dy\right),$$

$$\int_C R\,dz = \iint_S \left(-\frac{\partial R}{\partial x}\,dz \wedge dx + \frac{\partial R}{\partial y}\,dy \wedge dz\right).$$

La suma de esas tres ecuaciones nos da la fórmula (12.27) del teorema de Stokes. Puesto que las tres son parecidas, tan sólo demostramos la (12.28).

El plan de la demostración consiste en expresar la integral de superficie como una integral doble sobre T. Entonces se aplica el teorema de Green para expresar la integral doble sobre T como una integral de línea sobre Γ. Por último, demostramos que esta integral de línea es igual a $\int_C P\,dx$.

Escribamos

$$\boldsymbol{r}(u, v) = X(u, v)\boldsymbol{i} + Y(u, v)\boldsymbol{j} + Z(u, v)\boldsymbol{k}$$

y expresemos la integral de superficie sobre S en la forma

$$\iint_S \left(-\frac{\partial P}{\partial y}\,dx \wedge dy + \frac{\partial P}{\partial z}\,dz \wedge dx\right) = \iint_T \left\{-\frac{\partial P}{\partial y}\frac{\partial(X, Y)}{\partial(u, v)} + \frac{\partial P}{\partial z}\frac{\partial(Z, X)}{\partial(u, v)}\right\} du\,dv.$$

Designemos ahora con p la función compuesta dada por

$$p(u, v) = P[X(u, v), Y(u, v), Z(u, v)].$$

El último integrando puede escribirse en la forma

$$(12.29) \qquad -\frac{\partial P}{\partial y}\frac{\partial(X, Y)}{\partial(u, v)} + \frac{\partial P}{\partial z}\frac{\partial(Z, X)}{\partial(u, v)} = \frac{\partial}{\partial u}\left(p\,\frac{\partial X}{\partial v}\right) - \frac{\partial}{\partial v}\left(p\,\frac{\partial X}{\partial u}\right).$$

En el ejercicio 13 de la sección 12.13 se dan las directrices de la comprobación de (12.29). Aplicando a la integral doble sobre T el teorema de Green obtenemos

$$\iint_T \left\{\frac{\partial}{\partial u}\left(p\,\frac{\partial X}{\partial v}\right) - \frac{\partial}{\partial v}\left(p\,\frac{\partial X}{\partial u}\right)\right\} du\,dv = \int_\Gamma p\,\frac{\partial X}{\partial u}\,du + p\,\frac{\partial X}{\partial v}\,dv,$$

en donde Γ se recorre en sentido positivo. Parametricemos Γ mediante una función $\boldsymbol{\gamma}$ definida en un intervalo $[a, b]$ y sea

$$(12.30) \qquad \boldsymbol{\alpha}(t) = \boldsymbol{r}[\boldsymbol{\gamma}(t)]$$

la correspondiente parametrización de C. Expresando entonces cada integral de línea en función de su representación paramétrica encontramos que

$$\int_{\Gamma} p \frac{\partial X}{\partial u} du + p \frac{\partial X}{\partial v} dv = \int_C P\, dx,$$

lo cual completa la demostración de (12.28).

12.12 El rotacional y la divergencia de un campo vectorial

La integral de superficie que aparece en el teorema de Stokes puede expresarse en forma más sencilla en función del *rotacional* de un campo vectorial. Sea $\boldsymbol{F}$ un campo vectorial derivable dado por

$$\boldsymbol{F}(x, y, z) = P(x, y, z)\boldsymbol{i} + Q(x, y, z)\boldsymbol{j} + R(x, y, z)\boldsymbol{k}.$$

El rotacional de $\boldsymbol{F}$ es otro campo vectorial definido mediante la ecuación

$$\text{(12.31)} \qquad \operatorname{rot} \boldsymbol{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\boldsymbol{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\boldsymbol{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\boldsymbol{k}.$$

Los componentes de rot $\boldsymbol{F}$ son las funciones que aparecen en la integral de superficie de la fórmula de Stokes (12.27). Por consiguiente, esta integral de superficie puede ponerse en la forma

$$\iint_S (\operatorname{rot} \boldsymbol{F}) \cdot \boldsymbol{n}\, dS,$$

siendo $\boldsymbol{n}$ el vector unitario normal que tiene el mismo sentido que el producto vectorial fundamental de la superficie; esto es,

$$\boldsymbol{n} = \frac{\dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v}}{\left\| \dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v} \right\|}.$$

La integral de línea en la fórmula de Stokes (12.27) puede ponerse en la forma $\int_C \boldsymbol{F} \cdot d\boldsymbol{\alpha}$, siendo $\boldsymbol{\alpha}$ la representación de C dada por (12.30). Así pues, el teorema de Stokes toma la forma más sencilla

$$\iint_S (\operatorname{rot} \boldsymbol{F}) \cdot \boldsymbol{n}\, dS = \int_C \boldsymbol{F} \cdot d\boldsymbol{\alpha}.$$

Para el caso especial en el que S es una región del plano xy y $\boldsymbol{n} = \boldsymbol{k}$, esta fórmula se reduce a

$$\iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \int_C P\, dx + Q\, dy .$$

La ecuación (12.31) que define el rotacional puede recordarse fácilmente escribiéndola como un desarrollo de un determinante de tercer orden,

$$\operatorname{rot} \boldsymbol{F} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\boldsymbol{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\boldsymbol{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\boldsymbol{k} .$$

Este determinante está desarrollado por la primera fila, pero cada «producto» como el de $\partial/\partial y$ por R debe interpretarse como la derivada parcial $\partial R/\partial y$. También podemos escribir esa fórmula como un producto vectorial

$$\operatorname{rot} \boldsymbol{F} = \nabla \times \boldsymbol{F},$$

en donde el símbolo ∇ se maneja como si fuera un vector,

$$\nabla = \frac{\partial}{\partial x}\boldsymbol{i} + \frac{\partial}{\partial y}\boldsymbol{j} + \frac{\partial}{\partial z}\boldsymbol{k} .$$

Si formamos el «producto escalar» $\nabla \cdot \boldsymbol{F}$ de modo puramente formal, interpretando otra vez los productos tales como $\partial/\partial x$ por P como $\partial P/\partial x$, encontramos que

(12.32) $$\nabla \cdot \boldsymbol{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} .$$

La ecuación (12.32) define un campo escalar llamado *divergencia* de $\boldsymbol{F}$, y se escribe div $\boldsymbol{F}$.

Ya hemos utilizado el símbolo $\nabla\varphi$ para designar el gradiente de un campo escalar φ, dado por

$$\nabla\varphi = \frac{\partial \varphi}{\partial x}\boldsymbol{i} + \frac{\partial \varphi}{\partial y}\boldsymbol{j} + \frac{\partial \varphi}{\partial z}\boldsymbol{k} .$$

Esta fórmula puede interpretarse como una multiplicación formal del vector simbólico ∇ por el campo escalar φ. Así pues, el gradiente, la divergencia y el rotacional pueden representarse simbólicamente por los tres productos $\nabla\varphi$, $\nabla \cdot \boldsymbol{F}$, y $\nabla \times \boldsymbol{F}$, respectivamente.

Algunos de los teoremas antes demostrados pueden expresarse en función del rotacional. Por ejemplo, en el teorema 10.9 se demostró que un campo vectorial $\boldsymbol{f} = (f_1, \ldots, f_n)$, derivable con continuidad en un conjunto convexo abierto S del n-espacio, es un gradiente en S si y sólo si las derivadas parciales de los componentes de $\boldsymbol{f}$ satisfacen las relaciones

$$D_k f_j(\boldsymbol{x}) = D_j f_k(\boldsymbol{x}) \qquad (j, k = 1, 2, \ldots, n). \tag{12.33}$$

En el caso tri-dimensional, el teorema 10.9 puede enunciarse así.

TEOREMA 12.4. *Si $\boldsymbol{F} = P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k}$ es un campo vectorial derivable con continuidad en un conjunto convexo abierto S del espacio de 3 dimensiones, entonces $\boldsymbol{F}$ es un gradiente en S si y sólo si tenemos*

$$\text{rot}\, \boldsymbol{F} = \boldsymbol{O} \quad \textit{en} \quad S. \tag{12.34}$$

Demostración. En el caso tri-dimensional las relaciones (12.33) son equivalentes a la afirmación de que rot $\boldsymbol{F} = \boldsymbol{O}$.

12.13 Ejercicios

En cada ejercicio del 1 al 4, transformar la integral de superficie $\iint_S (\text{rot}\ \boldsymbol{F}) \cdot \boldsymbol{n}\, dS$ en una integral de línea utilizando el teorema de Stokes, y calcular entonces la integral de línea.

1. $\boldsymbol{F}(x, y, z) = y^2\boldsymbol{i} + xy\boldsymbol{j} + xz\boldsymbol{k}$, donde S es el hemisferio $x^2 + y^2 + z^2 = 1$, $z \geq 0$, y $\boldsymbol{n}$ es la normal unitaria con componente z no negativa.
2. $\boldsymbol{F}(x, y, z) = y\boldsymbol{i} + z\boldsymbol{j} + x\boldsymbol{k}$, donde S es la parte de paraboloide $z = 1 - x^2 - y^2$ con $z \geq 0$, y $\boldsymbol{n}$ es la normal unitaria con componente z no negativa.
3. $\boldsymbol{F}(x, y, z) = (y - z)\boldsymbol{i} + yz\boldsymbol{j} - xz\boldsymbol{k}$, donde S consta de las cinco caras del cubo $0 \leq x \leq 2$, $0 \leq y \leq 2$, $0 \leq z \leq 2$ no situadas en el plano xy. $\boldsymbol{n}$ es la normal unitaria exterior.
4. $\boldsymbol{F}(x, y, z) = xz\boldsymbol{i} - y\boldsymbol{i} + x^2y\boldsymbol{k}$, donde S consta de las tres caras no situadas en el plano xz del tetraedro limitado por los tres planos coordenados y el plano $3x + y + 3z = 6$. La normal $\boldsymbol{n}$ es la normal unitaria exterior del tetraedro.

En los ejercicios del 5 al 10, usar el teorema de Stokes para demostrar que las integrales de línea tienen los valores que se dan. En cada caso, explicar el sentido en el que se recorre C para llegar al resultado.

5. $\int_C y\, dx + z\, dy + x\, dz = \pi a^2\sqrt{3}$, siendo C la curva de intersección de la esfera $x^2 + y^2 + z^2 = a^2$ y el plano $x + y + z = 0$.

6. $\int_C (y + z)\,dx + (z + x)\,dy + (x + y)\,dz = 0$, siendo C la curva de intersección del cilindro $x^2 + y^2 = 2y$ y el plano $y = z$.
7. $\int_C y^2\,dx + xy\,dy + xz\,dz = 0$, donde C es la curva del ejercicio 6.
8. $\int_C (y - z)\,dx + (z - x)\,dy + (x - y)\,dz = 2\pi a(a + b)$, donde C es la intersección del cilindro $x^2 + y^2 = a^2$ y el plano $x/a + z/b = 1$, $a > 0$, $b > 0$.
9. $\int_C (y^2 + z^2)\,dx + (x^2 + z^2)\,dy + (x^2 + y^2)\,dz = 2\pi ab^2$, siendo C la intersección del hemisferio $x^2 + y^2 + z^2 = 2ax$, $z > 0$, y el cilindro $x^2 + y^2 = 2bx$, donde $0 < b < a$.
10. $\int_C (y^2 - z^2)\,dx + (z^2 - x^2)\,dy + (x^2 - y^2)\,dz = 9a^3/2$, donde C es la curva intersección de la superficie del cubo $0 \le x \le a, 0 \le y \le a, 0 \le z \le a$ y el plano $x + y + z = 3a/2$.
11. Si $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ y $P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k} = \boldsymbol{a} \times \boldsymbol{r}$, siendo $\boldsymbol{a}$ un vector constante, demostrar que $\int_C P\,dx + Q\,dy + R\,dz = 2\iint_S \boldsymbol{a} \cdot \boldsymbol{n}\,dS$, donde C es una curva que limita una superficie paramétrica S y $\boldsymbol{n}$ es la normal a S adecuada.
12. Sea $\boldsymbol{F} = P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k}$, donde $P = -y/(x^2 + y^2)$, $Q = x/(x^2 + y^2)$, $R = z$. Sea D el toro engendrado por la rotación de la circunferencia $(x - 2)^2 + z^2 = 1$, $y = 0$, alrededor del eje z. Demostrar que $\operatorname{rot} \boldsymbol{F} = \boldsymbol{O}$ pero que $\int_C P\,dx + Q\,dy + R\,dz$ no es cero si la curva C es la circunferencia $x^2 + y^2 = 4$, $z = 0$.
13. Este ejercicio da indicaciones para una demostración de la ecuación (12.29) usada al demostrar el teorema de Stokes.

a) Utilizar la fórmula de la derivación de un producto para probar que

$$\frac{\partial}{\partial u}\left(p\,\frac{\partial X}{\partial v}\right) - \frac{\partial}{\partial v}\left(p\,\frac{\partial X}{\partial u}\right) = \frac{\partial p}{\partial u}\,\frac{\partial X}{\partial v} - \frac{\partial p}{\partial v}\,\frac{\partial X}{\partial u}\,.$$

b) Póngase $p(u, v) = P[X(u, v), Y(u, v), Z(u, v)]$. Calcular $\partial p/\partial u$ y $\partial p/\partial v$ por medio de la regla de la cadena y aplicar la parte a) para deducir la ecuación (12.29),

$$\frac{\partial}{\partial u}\left(p\,\frac{\partial X}{\partial v}\right) - \frac{\partial}{\partial v}\left(p\,\frac{\partial X}{\partial u}\right) = -\frac{\partial P}{\partial y}\,\frac{\partial(X, Y)}{\partial(u, v)} + \frac{\partial P}{\partial z}\,\frac{\partial(Z, X)}{\partial(u, v)}\,.$$

12.14 Otras propiedades del rotacional y de la divergencia

El rotacional y la divergencia de un campo vectorial están relacionados con la matriz jacobiana. Si $\boldsymbol{F} = P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k}$, la matriz jacobiana de $\boldsymbol{F}$ es

$$DF(x, y, z) = \begin{bmatrix} \dfrac{\partial P}{\partial x} & \dfrac{\partial P}{\partial y} & \dfrac{\partial P}{\partial z} \\[2ex] \dfrac{\partial Q}{\partial x} & \dfrac{\partial Q}{\partial y} & \dfrac{\partial Q}{\partial z} \\[2ex] \dfrac{\partial R}{\partial x} & \dfrac{\partial R}{\partial y} & \dfrac{\partial R}{\partial z} \end{bmatrix}.$$

La traza de esta matriz (la suma de elementos de su diagonal principal) es la divergencia de $\boldsymbol{F}$.

Toda matriz real A puede escribirse como suma de una matriz simétrica, $\frac{1}{2}(A + A^t)$ y de una matriz hemi-simétrica, $\frac{1}{2}(A - A^t)$. Cuando A es la matriz jacobiana DF, la parte hemi-simétrica se convierte en

$$\frac{1}{2}\begin{bmatrix} 0 & \dfrac{\partial P}{\partial y} - \dfrac{\partial Q}{\partial x} & \dfrac{\partial P}{\partial z} - \dfrac{\partial R}{\partial x} \\ \dfrac{\partial Q}{\partial x} - \dfrac{\partial P}{\partial y} & 0 & \dfrac{\partial Q}{\partial z} - \dfrac{\partial R}{\partial y} \\ \dfrac{\partial R}{\partial x} - \dfrac{\partial P}{\partial z} & \dfrac{\partial R}{\partial y} - \dfrac{\partial Q}{\partial z} & 0 \end{bmatrix}. \tag{12.35}$$

Los elementos no nulos de esta matriz son los componentes de rot $\boldsymbol{F}$ y sus opuestos. Si la matriz jacobiana DF es simétrica, cada elemento de (12.35) es cero y rot $\boldsymbol{F} = \boldsymbol{O}$.

EJEMPLO 1. Sea $\boldsymbol{F}(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$. Tenemos entonces

$$P(x, y, z) = x, \qquad Q(x, y, z) = y, \qquad R(x, y, z) = z,$$

y la matriz jacobiana correspondiente es la matriz idéntica de orden 3×3. Por lo tanto

$$\text{div } \boldsymbol{F} = 3 \qquad \text{y} \qquad \text{rot } \boldsymbol{F} = \boldsymbol{O}.$$

En general, si $\boldsymbol{F}(x, y, z) = f(x)\boldsymbol{i} + g(y)\boldsymbol{j} + h(z)\boldsymbol{k}$, la matriz jacobiana tiene los elementos $f'(x)$, $g'(y)$, $h'(z)$ en la diagonal principal y ceros los restantes elementos, así que

$$\text{div } \boldsymbol{F} = f'(x) + g'(y) + h'(z) \qquad \text{y} \qquad \text{rot } \boldsymbol{F} = \boldsymbol{O}.$$

EJEMPLO 2. Sea $\boldsymbol{F}(x, y, z) = xy^2z^2\boldsymbol{i} + z^2 \operatorname{sen} y\, \boldsymbol{j} + x^2e^y\boldsymbol{k}$. La matriz jacobiana es

$$\begin{bmatrix} y^2z^2 & 2xyz^2 & 2xy^2z \\ 0 & z^2 \cos y & 2z \operatorname{sen} y \\ 2xe^y & x^2e^y & 0 \end{bmatrix}.$$

Por lo tanto,

$$\text{div } \boldsymbol{F} = y^2z^2 + z^2 \cos y$$

y

$$\text{rot } \boldsymbol{F} = (x^2e^y - 2z \operatorname{sen} y)\boldsymbol{i} + (2xy^2z - 2xe^y)\boldsymbol{j} - 2xyz^2\boldsymbol{k}.$$

EJEMPLO 3. *Divergencia y rotacional de un gradiente.* Supongamos que F es un gradiente, sea $F = \text{grad } \varphi = \partial\varphi/\partial x\, i + \partial\varphi/\partial y\, j + \partial\varphi/\partial z\, k$. La matriz jacobiana es

$$(12.36) \qquad \begin{bmatrix} \dfrac{\partial^2\varphi}{\partial x^2} & \dfrac{\partial^2\varphi}{\partial y\,\partial x} & \dfrac{\partial^2\varphi}{\partial z\,\partial x} \\ \dfrac{\partial^2\varphi}{\partial x\,\partial y} & \dfrac{\partial^2\varphi}{\partial y^2} & \dfrac{\partial^2\varphi}{\partial z\,\partial y} \\ \dfrac{\partial^2\varphi}{\partial x\,\partial z} & \dfrac{\partial^2\varphi}{\partial y\,\partial z} & \dfrac{\partial^2\varphi}{\partial z^2} \end{bmatrix}.$$

Por consiguiente,

$$\text{div } F = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2}.$$

La expresión del segundo miembro se llama *laplaciana* de φ y a menudo se representa con el símbolo $\nabla^2\varphi$. Así pues, la divergencia de un gradiente $\nabla\varphi$ es la laplaciana de φ. Simbólicamente esto se expresa poniendo

$$(12.37) \qquad \text{div } (\nabla\varphi) = \nabla^2\varphi.$$

Cuando $\nabla^2\varphi = 0$, la función φ se llama *armónica.* La ecuación (12.37) hace ver que el gradiente de una función armónica tiene divergencia nula. Cuando las derivadas parciales mixtas que aparecen en la matriz (12.36) son continuas, la matriz es simétrica y rot F es cero. Es decir,

$$\text{rot } (\text{grad } \varphi) = O$$

para todo campo escalar φ con derivadas parciales segundas mixtas continuas. Este ejemplo demuestra que rot $F = O$ es condición necesaria para que un campo vectorial F derivable con continuidad sea un gradiente. Dicho de otro modo, si rot $F \neq O$ en un conjunto abierto S, entonces F no es un gradiente en S. Sabemos también, según el teorema 12.4 que si rot $F = O$ en un conjunto *convexo* abierto S, entonces F es un gradiente en S. Un campo con rotacional nulo se llama *irrotacional.*

EJEMPLO 4. *Campo vectorial con divergencia y rotacional nulos.* Sea S el conjunto de todos los puntos $(x, y) \neq (0, 0)$, y pongamos

$$F(x, y) = -\frac{y}{x^2 + y^2}\, i + \frac{x}{x^2 + y^2}\, j$$

si $(x, y) \in S$. Según el ejemplo 2 de la sección 10.16 sabemos que $\boldsymbol{F}$ *no es un gradiente de* S (si bien $\boldsymbol{F}$ es un gradiente en todo rectángulo que no contenga el origen). La matriz jacobiana es

$$DF(x, y) = \begin{bmatrix} \dfrac{2xy}{(x^2+y^2)^2} & \dfrac{y^2-x^2}{(x^2+y^2)^2} & 0 \\ \dfrac{y^2-x^2}{(x^2+y^2)^2} & \dfrac{-2xy}{(x^2+y^2)^2} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

y vemos al momento que div $\boldsymbol{F} = 0$ y rot $\boldsymbol{F} = \boldsymbol{O}$ en S.

EJEMPLO 5. *Divergencia y rotacional de un rotacional. Si* $\boldsymbol{F} = P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k}$, el rotacional de $\boldsymbol{F}$ es un nuevo campo vectorial y podemos calcular su divergencia y su rotacional. La matriz jacobiana de rot $\boldsymbol{F}$ es

$$\begin{bmatrix} \dfrac{\partial^2 R}{\partial x\,\partial y} - \dfrac{\partial^2 Q}{\partial x\,\partial z} & \dfrac{\partial^2 R}{\partial y^2} - \dfrac{\partial^2 Q}{\partial y\,\partial z} & \dfrac{\partial^2 R}{\partial z\,\partial y} - \dfrac{\partial^2 Q}{\partial z^2} \\ \dfrac{\partial^2 P}{\partial x\,\partial z} - \dfrac{\partial^2 R}{\partial x^2} & \dfrac{\partial^2 P}{\partial y\,\partial z} - \dfrac{\partial^2 R}{\partial y\,\partial x} & \dfrac{\partial^2 P}{\partial z^2} - \dfrac{\partial^2 R}{\partial z\,\partial x} \\ \dfrac{\partial^2 Q}{\partial x^2} - \dfrac{\partial^2 P}{\partial x\,\partial y} & \dfrac{\partial^2 Q}{\partial y\,\partial x} - \dfrac{\partial^2 P}{\partial y^2} & \dfrac{\partial^2 Q}{\partial z\,\partial x} - \dfrac{\partial^2 P}{\partial z\,\partial y} \end{bmatrix}.$$

Si suponemos que todas las derivadas parciales mixtas son continuas, encontramos que

$$\text{div}\,(\text{rot}\,\boldsymbol{F}) = 0$$

y

(12.38) $$\text{rot}\,(\text{rot}\,\boldsymbol{F}) = \text{grad}\,(\text{div}\,\boldsymbol{F}) - \nabla^2\boldsymbol{F},$$

estando definida $\nabla^2 \boldsymbol{F}$ mediante la ecuación

$$\nabla^2\boldsymbol{F} = (\nabla^2 P)\boldsymbol{i} + (\nabla^2 Q)\boldsymbol{j} + (\nabla^2 R)\boldsymbol{k}.$$

La identidad (12.38) relaciona los cuatro operadores, gradiente, rotacional, divergencia y laplaciana. En el ejercicio 7 de la sección 12.15 se propone la comprobación de (12.38).

El rotacional y la divergencia tienen algunas propiedades generales análogas

a las de las derivadas ordinarias. En primer lugar, son *operadores lineales*. Esto es, si a y b son constantes, tenemos

(12.39) $$\operatorname{div}(a\boldsymbol{F} + b\boldsymbol{G}) = a \operatorname{div} \boldsymbol{F} + b \operatorname{div} \boldsymbol{G},$$

y

(12.40) $$\operatorname{rot}(a\boldsymbol{F} + b\boldsymbol{G}) = a \operatorname{rot} \boldsymbol{F} + b \operatorname{rot} \boldsymbol{G}.$$

También poseen la propiedad análoga a la fórmula de la derivación de un producto:

(12.41) $$\operatorname{div}(\varphi\boldsymbol{F}) = \varphi \operatorname{div} \boldsymbol{F} + \nabla\varphi \cdot \boldsymbol{F},$$

y

(12.42) $$\operatorname{rot}(\varphi\boldsymbol{F}) = \varphi \operatorname{rot} \boldsymbol{F} + \nabla\varphi \times \boldsymbol{F},$$

donde φ es cualquier campo escalar derivable. Estas propiedades son consecuencias inmediatas de las definiciones de rotacional y divergencia; sus demostraciones se proponen en el ejercicio 6 de la sección 12.15.

Si empleamos el vector simbólico

$$\nabla = \frac{\partial}{\partial x}\boldsymbol{i} + \frac{\partial}{\partial y}\boldsymbol{j} + \frac{\partial}{\partial z}\boldsymbol{k}$$

una vez más, cada una de las fórmulas (12.41) y (12.42) toma una forma muy parecida a la regla de la derivación de un producto:

$$\nabla \cdot (\varphi\boldsymbol{F}) = \varphi \nabla \cdot \boldsymbol{F} + \nabla\varphi \cdot \boldsymbol{F}$$

y

$$\nabla \times (\varphi\boldsymbol{F}) = \varphi \nabla \times \boldsymbol{F} + \nabla\varphi \times \boldsymbol{F}.$$

En el ejemplo 3 la laplaciana de un campo escalar, $\nabla^2\varphi$, se definió como $\partial^2\varphi/\partial x^2 + \partial^2\varphi/\partial y^2 + \partial^2\varphi/\partial z^2$. En el ejemplo 5 la laplaciana $\nabla^2\boldsymbol{F}$ de un campo vectorial se definió por medio de los componentes. Obtenemos fórmulas correctas para $\nabla^2\varphi$ y para $\nabla^2\boldsymbol{F}$ si interpretamos ∇^2 como el operador simbólico

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Esta fórmula para ∇^2 también resulta por multiplicación interior del vector simbólico ∇ por sí mismo.
Así, tenemos $\nabla^2 = \nabla \cdot \nabla$ y podemos escribir:

$$\nabla^2\varphi = (\nabla \cdot \nabla)\varphi \qquad \text{y} \qquad \nabla^2 \boldsymbol{F} = (\nabla \cdot \nabla)\boldsymbol{F}.$$

Consideremos ahora la fórmula $\nabla \cdot \nabla\varphi$. Ésta puede interpretarse como $(\nabla \cdot \nabla)\varphi$, que es $\nabla^2\varphi$; o como $\nabla \cdot (\nabla\varphi)$, que es div $(\nabla\varphi)$. En el ejemplo 3 se demostró que div $(\nabla\varphi) = \nabla^2\varphi$, así tenemos,

$$(\nabla \cdot \nabla)\varphi = \nabla \cdot (\nabla\varphi);$$

luego podemos escribir $\nabla \cdot \nabla\varphi$ para una u otra de esas expresiones sin peligro de ambigüedad. Esto no es cierto, sin embargo, cuando φ se sustituye por un campo vectorial $\boldsymbol{F}$. La expresión $(\nabla \cdot \nabla)\boldsymbol{F}$ es $\nabla^2\boldsymbol{F}$, que ha sido definida. No obstante, $\nabla \cdot (\nabla \boldsymbol{F})$ no tiene significado a causa de que $\nabla \boldsymbol{F}$ no está definido. Por lo tanto, la expresión $\nabla \cdot \nabla \boldsymbol{F}$ tiene significado únicamente cuando se interpreta como $(\nabla \cdot \nabla)\boldsymbol{F}$. Estas observaciones hacen ver que si bien las fórmulas simbólicas sirven algunas veces como notación manejable y mnemotécnica, es preciso utilizarlas con mucha precaución.

12.15 Ejercicios

1. Para cada uno de los siguientes campos vectoriales determinar la matriz jacobiana y calcular el rotacional y la divergencia
   a) $\boldsymbol{F}(x, y, z) = (x^2 + yz)\boldsymbol{i} + (y^2 + xz)\boldsymbol{j} + (z^2 + xy)\boldsymbol{k}$.
   b) $\boldsymbol{F}(x, y, z) = (2z - 3y)\boldsymbol{i} + (3x - z)\boldsymbol{j} + (y - 2x)\boldsymbol{k}$.
   c) $\boldsymbol{F}(x, y, z) = (z + \operatorname{sen} y)\boldsymbol{i} - (z - x\cos y)\boldsymbol{j}$.
   d) $\boldsymbol{F}(x, y, z) = e^{xy}\boldsymbol{i} + \cos xy\,\boldsymbol{j} + \cos xz^2\boldsymbol{k}$.
   e) $\boldsymbol{F}(x, y, z) = x^2 \operatorname{sen} y\,\boldsymbol{i} + y^2 \operatorname{sen} xz\,\boldsymbol{j} + xy \operatorname{sen}(\cos z)\boldsymbol{k}$.
2. Si $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ y $r = \|\boldsymbol{r}\|$, calcular rot $[f(r)\boldsymbol{r}]$, siendo f una función derivable.
3. Si $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ y $\boldsymbol{A}$ es un vector constante, demostrar que $\operatorname{rot}(\boldsymbol{A} \times \boldsymbol{r}) = 2\boldsymbol{A}$.
4. Si $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ y $r = \|\boldsymbol{r}\|$, hallar todos los valores de n para los que $(r^n\boldsymbol{r}) = 0$.
5. Hallar un campo vectorial cuyo rotacional es $x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ o demostrar que no existe un tal campo vectorial.
6. Demostrar las propiedades elementales del rotacional y de la divergencia expresada en las ecuaciones de la (12.39) a la (12.42).
7. Demostrar que $\operatorname{rot}(\operatorname{rot} \boldsymbol{F}) = \operatorname{grad}(\operatorname{div} \boldsymbol{F}) - \nabla^2\boldsymbol{F}$ si los componentes de $\boldsymbol{F}$ tienen derivadas parciales mixtas de segundo orden continuas.
8. Demostrar la identidad

$$\nabla \cdot (\boldsymbol{F} \times \boldsymbol{G}) = \boldsymbol{G} \cdot (\nabla \times \boldsymbol{F}) - \boldsymbol{F} \cdot (\nabla \times \boldsymbol{G}),$$

   donde $\boldsymbol{F}$ y $\boldsymbol{G}$ son campos vectoriales diferenciables.

9. Un campo vectorial F no será gradiente de un potencial a menos que rot $F = O$. No obstante, es posible encontrar un campo escalar no nulo μ tal que μF es un gradiente. Demostrar que si un tal μ existe, F es siempre perpendicular a su rotacional. Cuando el campo es bidimensional, $F = Pi + Qj$, este ejercicio nos da una condición necesaria para que la ecuación diferencial $P\,dx + Q\,dy = 0$ posea un factor integrante. (El recíproco también es cierto. Esto es, si $F \cdot \text{rot}\, F = 0$ en una región conveniente, existe un μ no nulo tal que μF es un gradiente. La demostración del recíproco no se pide.)
10. Sea $F(x, y, z) = y^2z^2 i + z^2x^2 j + x^2y^2 k$. Demostrar que rot F no es siempre cero, pero que $F \cdot \text{rot}\, F = 0$. Hallar un campo escalar μ tal que μF sea un gradiente.
11. Sea $V(x, y) = y^c i + x^c j$, donde c es una constante positiva, y sea $r(x, y) = xi + yj$. Consideremos una región plana R bordeada por una curva de Jordan regular a trozos C. Calcular div $(V \times r)$ y rot $(V \times r)$ y aplicando el teorema de Green demostrar que

$$\oint_C V \times r \cdot d\alpha = 0,$$

donde α es la función que describe C.
12. Demostrar que el teorema de Green puede expresarse en la forma

$$\iint_R (\text{rot}\ V) \cdot k\, dx\, dy = \oint_C V \cdot T\, ds,$$

donde T es el vector unitario tangente a C y s es la longitud de arco.
13. Una región plana R está limitada por una curva de Jordan regular a trozos C. Se conocen los momentos de inercia de R alrededor de los ejes x e y que valen respectivamente a y b. Calcular la integral de línea

$$\oint_C \nabla(r^4) \cdot n\, ds$$

en función de a y b. En esta integral, $r = \|xi + yj\|$, n representa el vector unitario normal exterior a C y s es la longitud de arco. La curva se recorre en sentido contrario al de las agujas del reloj.
14. Sea F un campo vectorial bidimensional. Dar una definición de la integral de línea $\int_C F \times d\alpha$. Esa definición debe ser tal que pueda obtenerse como consecuencia del teorema de Green la fórmula siguiente:

$$\int_C F \times d\alpha = k \iint_R (\text{div}\, F)\, dx\, dy,$$

siendo R una región plana limitada por una curva cerrada simple C.

* 12.16 Reconstrucción de un campo vectorial a partir de su rotacional

Al estudiar el gradiente hemos aprendido a determinar si un campo vectorial dado es o no un gradiente. Consideramos ahora la cuestión análoga relativa al rota-

cional. Dado un campo vectorial $\boldsymbol{F}$, ¿hay un $\boldsymbol{G}$ tal que rot $\boldsymbol{G} = \boldsymbol{F}$? Supongamos que $\boldsymbol{F} = P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k}$ y $\boldsymbol{G} = L\boldsymbol{i} + M\boldsymbol{j} + N\boldsymbol{k}$. Para resolver la ecuación rot $\boldsymbol{G} = \boldsymbol{F}$ tenemos que resolver el sistema de ecuaciones diferenciales en derivadas parciales

$$\text{(12.43)} \qquad \frac{\partial N}{\partial y} - \frac{\partial M}{\partial z} = P, \qquad \frac{\partial L}{\partial z} - \frac{\partial N}{\partial x} = Q, \qquad \frac{\partial M}{\partial x} - \frac{\partial L}{\partial y} = R$$

en las que P, Q y R son conocidas y L, M, N son las funciones incógnitas.

No siempre es posible resolver tal sistema. Por ejemplo, se demostró en la sección 12.14 que la divergencia de un rotacional es siempre cero. Por tanto, para que el sistema (12.43) tenga solución en un cierto conjunto abierto S es necesario que

$$\text{(12.44)} \qquad \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = 0$$

en todo S. Esta condición es también suficiente si restringimos convenientemente el conjunto S en el que se verifica (12.44). A continuación demostramos que la condición (12.44) es suficiente cuando S es un intervalo tridimensional.

TEOREMA 12.5. *Si $\boldsymbol{F}$ es derivable con continuidad en un intervalo abierto S de E_3, entonces existe un campo vectorial $\boldsymbol{G}$ tal que rot $\boldsymbol{G} = \boldsymbol{F}$ si y sólo si* div $\boldsymbol{F} = 0$ *en todo S.*

Demostración. La necesidad de la condición div $\boldsymbol{F} = 0$ ha sido ya establecida, puesto que la divergencia de un rotacional siempre es cero. Para establecer la suficiencia tenemos que encontrar tres funciones L, M, y N que satisfagan las tres ecuaciones (12.43). Intentemos resolver el problema tomando $L = 0$. Entonces la segunda y tercera ecuaciones (12.43) se transforman en

$$\frac{\partial N}{\partial x} = -Q \qquad \text{y} \qquad \frac{\partial M}{\partial x} = R.$$

Esto indica que debe ser:

$$N(x, y, z) = -\int_{x_0}^{x} Q(t, y, z)\, dt + f(y, z)$$

y

$$M(x, y, z) = \int_{x_0}^{x} R(t, y, z)\, dt + g(y, z),$$

donde cada integración se efectúa a lo largo de un segmento rectilíneo contenido en S y las «constantes de integración» $f(y, z)$ y $g(y, z)$ son independientes de x.

Para hallar una solución hagamos $f(y, z) = 0$. La primera ecuación (12.43) exige

$$\frac{\partial N}{\partial y} - \frac{\partial M}{\partial z} = P. \tag{12.45}$$

Según la elección de M y N que se ha indicado tenemos:

$$\frac{\partial N}{\partial y} - \frac{\partial M}{\partial z} = -\frac{\partial}{\partial y}\int_{x_0}^{x} Q(t, y, z)\,dt - \frac{\partial}{\partial z}\int_{x_0}^{x} R(t, y, z)\,dt - \frac{\partial g}{\partial z}. \tag{12.46}$$

Seguidamente permutamos las dos operaciones de derivación parcial e integración aplicando el teorema 10.8. Esto es, escribimos:

$$\frac{\partial}{\partial y}\int_{x_0}^{x} Q(t, y, z)\,dt = \int_{x_0}^{x} D_2Q(t, y, z)\,dt \tag{12.47}$$

y

$$\frac{\partial}{\partial z}\int_{x_0}^{x} R(t, y, z)\,dt = \int_{x_0}^{x} D_3R(t, y, z)\,dt. \tag{12.48}$$

Con ello la ecuación (12.46) se convierte en

$$\frac{\partial N}{\partial y} - \frac{\partial M}{\partial z} = \int_{x_0}^{x} [-D_2Q(t, y, z) - D_3R(t, y, z)]\,dt - \frac{\partial g}{\partial z}. \tag{12.49}$$

En virtud de la condición (12.44) podemos reemplazar el integrando de (12.49) por $D_1P(t, y, z)$; la ecuación (12.49) toma la forma

$$\frac{\partial N}{\partial y} - \frac{\partial M}{\partial z} = \int_{x_0}^{x} D_1P(t, y, z)\,dt - \frac{\partial g}{\partial z} = P(x, y, z) - P(x_0, y, z) - \frac{\partial g}{\partial z}.$$

Por consiguiente, (12.45) se satisfará si elegimos g de modo que $\partial g/\partial z = -P(x_0, y, z)$. Así, por ejemplo, podemos tomar

$$g(y, z) = -\int_{z_0}^{z} P(x_0, y, u)\,du.$$

Este razonamiento nos lleva a considerar el campo vectorial $G = Li + Mj + Nk$, donde $L(x, y, z) = 0$ y

$$M(x, y, z) = \int_{x_0}^{x} R(t, y, z)\,dt - \int_{z_0}^{z} P(x_0, y, u)\,du, \qquad N(x, y, z) = -\int_{x_0}^{x} Q(t, y, z)\,dt.$$

Con estas tres funciones L, M y N es fácil comprobar, por medio de (12.47) y (12.48), que se satisfacen las tres ecuaciones (12.43), obteniendo rot $\boldsymbol{G} = \boldsymbol{F}$, como deseábamos.

Debe observarse que la demostración precedente no sólo establece la existencia de un campo vectorial $\boldsymbol{G}$ cuyo rotacional es $\boldsymbol{F}$, sino que también proporciona un método directo para determinar $\boldsymbol{G}$ por medio de integrales calculadas con los componentes de $\boldsymbol{F}$.

Para una $\boldsymbol{F}$ dada, el campo vectorial $\boldsymbol{G}$ que hemos construido no es la única solución de la ecuación rot $\boldsymbol{G} = \boldsymbol{F}$. Si sumamos a $\boldsymbol{G}$ cualquier gradiente derivable con continuidad $\nabla \varphi$ obtenemos otra solución ya que

$$\text{rot } (\boldsymbol{G} + \nabla\varphi) = \text{ rot } \boldsymbol{G} + \text{ rot } (\nabla\varphi) = \text{ rot } \boldsymbol{G} = \boldsymbol{F},$$

puesto que rot $(\nabla \varphi) = \boldsymbol{O}$. Además, es fácil demostrar que *todas* las soluciones derivables con continuidad deben ser de la forma $\boldsymbol{G} + \nabla \varphi$. En realidad, si $\boldsymbol{H}$ es otra solución, entonces rot $\boldsymbol{H} =$ rot $\boldsymbol{G}$,de manera que rot $(\boldsymbol{H} - \boldsymbol{G}) = \boldsymbol{O}$. Según el teorema 10.9 resulta que $\boldsymbol{H} - \boldsymbol{G} = \nabla \varphi$ para algún gradiente derivable con continuidad $\nabla \varphi$; luego $\boldsymbol{H} = \boldsymbol{G} + \nabla\varphi$, como se afirmó.

Un campo vectorial $\boldsymbol{F}$ para el que div $\boldsymbol{F} = 0$ suele llamarse *solenoidal*. El teorema 12.5 establece que un campo vectorial es solenoidal en un paralelepípedo rectangular abierto S si, y sólo si, es el rotacional de otro campo vectorial en S

El ejemplo que sigue demuestra que esta afirmación no es válida para conjuntos abiertos *cualesquiera*.

EJEMPLO. *Un campo vectorial solenoidal que no es un rotacional.* Sea D la parte de E_3 comprendida entre dos esferas concéntricas de centro en el origen y radios a y b, siendo $0 < a < b$. Sea $\boldsymbol{V} = \boldsymbol{r}/r^3$, donde $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ y $r = \|\boldsymbol{r}\|$. Es fácil comprobar que div $\boldsymbol{V} = 0$ en todo D. En efecto, tenemos la fórmula general

$$\text{div } (r^n\boldsymbol{r}) = (n + 3)r^n,$$

y en este ejemplo $n = -3$. Usaremos el teorema de Stokes para demostrar que $\boldsymbol{V}$ *no es* un rotacional en D (si bien lo es en todo intervalo tridimensional abierto que no contenga el origen). Para ello supongamos que existe un campo vectorial $\boldsymbol{U}$ tal que $\boldsymbol{V} =$ rot $\boldsymbol{U}$ en D y llegamos a una contradicción. Según el teorema de Stokes podemos escribir

$$\iint_S (\text{rot } \boldsymbol{U}) \cdot \boldsymbol{n}\, dS = \oint_C \boldsymbol{U} \cdot d\boldsymbol{\alpha}, \tag{12.50}$$

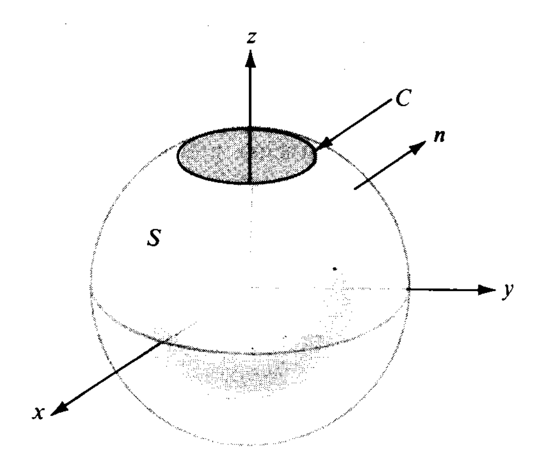


FIGURA 12.13 *La superficie S y la curva C de la ecuación* (12.50).

donde S y C son la superfice y la curva dibujada en la figura 12.13. Para construir S, tomamos una superficie esférica de radio R concéntrica con las fronteras de D, siendo $a < R < b$, y quitamos un pequeño casquete polar como se indica en la figura. La parte que queda es la superficie S. La curva C es el borde circular dibujado. Representemos con $\boldsymbol{n}$ la normal unitaria exterior a S, de modo que $\boldsymbol{n} = \boldsymbol{r}/r$. Puesto que rot $\boldsymbol{U} = \boldsymbol{V} = \boldsymbol{r}/r^3$, tenemos

$$(\text{rot } \boldsymbol{U}) \cdot \boldsymbol{n} = \frac{\boldsymbol{r}}{r^3} \cdot \frac{\boldsymbol{r}}{r} = \frac{1}{r^2}.$$

Sobre la superficie S este producto vectorial tiene el valor constante $1/R^2$. Por lo tanto, tenemos,

$$\iint_S (\text{rot } \boldsymbol{U}) \cdot \boldsymbol{n} \, dS = \frac{1}{R^2} \iint_S dS = \frac{\text{área de } S}{R^2}.$$

Cuando el casquete polar degenera en un punto el área de S tiende a $4\pi R^2$ (área de toda la esfera) y por consiguiente, el valor de la integral de superficie (12.50) tiende a 4π.

Examinemos seguidamente la integral de línea de (12.50). Es fácil demostrar que para toda integral de línea $\int_C \boldsymbol{U} \cdot d\boldsymbol{\alpha}$ tenemos la desigualdad

$$\left| \int_C \boldsymbol{U} \cdot d\boldsymbol{\alpha} \right| \leq M \cdot (\text{longitud de } C)$$

donde M es una constante que depende de $\boldsymbol{U}$. (En efecto, M puede tomarse como

el máximo de $\|U\|$ en C). Por consiguiente, al reducir el casquete polar a un punto, la longitud de C y el valor de la integral de línea tienden ambos a cero. Llegamos así a una contradicción; la integral de superficie (12.50) tiende a 4π, y la correspondiente integral de línea a la que es igual tiende a 0. Luego no puede existir en la región D una función U cuyo rotacional sea V.

La dificultad aquí es debida a la estructura geométrica de la región D. Si bien esta región es simplemente conexa (esto es, toda curva cerrada simple en D es el borde de una superficie paramétrica contenida totalmente en D) existen *superficies* cerradas en D que no son frontera completa de un sólido contenido enteramente en D. Por ejemplo, ninguna esfera que incluya el origen es la frontera completa de un sólido enteramente contenido en D. Si la región D tiene la propiedad de que *toda* superficie cerrada contenida en D es la frontera de un sólido enteramente contenido en D, puede demostrarse que existe un campo vectorial U tal que $V = \text{rot } U$ en D si, y sólo si, div $V = 0$ en toda región D. La demostración es difícil y no la daremos aquí.

* 12.17 Ejercicios

1. Hallar un campo vectorial $G(x, y, z)$ cuyo rotacional es $2i + j + 3k$ en todo E_3. ¿Cuál es el campo vectorial derivable con continuidad del tipo más general con esa propiedad?
2. Demostrar que el campo vectorial $F(x, y, z) = (y - z)i + (z - x)j + (x - y)k$ es solenoidal, y hallar un campo vectorial G tal que $F = \text{rot } G$
3. Sea $F(x, y, z) = -zi + xyk$. Hallar un campo vectorial derivable con continuidad G de la forma $G(x,y,z)=L(x,y,z)i+M(x,y,z)j$ tal que $F = \text{rot } G$ en todo paralelepípedo rectangular de E_3. ¿Cuál es el G más general de esa forma?
4. Si dos campos vectoriales U y V son irrotacionales, demostrar que el campo vectorial $U \times V$ es solenoidal.
5. Sea $r = xi + yj + zk$ y $r = \|r\|$. Demostrar que $n = -3$ es el único valor de n para el que $r^n r$ es solenoidal siendo $r \neq 0$. Para este n, elegir un paralelepípedo S que no contenga el origen y expresar $r^{-3}r$ como un rot en S. [*Advertencia.* Si bien $r^{-3}r$ es un rotacional en cualquier S de este tipo, *no lo es* en el conjunto de todos los puntos distintos del $(0, 0, 0)$].
6. Encontrar la forma más general de una función f derivable con continuidad y de una sola variable real tal que el campo vectorial $f(r)r$ sea solenoidal, donde $r = xi + yj + zk$ y $r = \|r\|$.
7. Sea V un campo vectorial derivable con continuidad en cierto paralelepípedo rectangular S de E_3. Consideremos las dos afirmaciones siguientes relativas a V:
   (I) rot $V = O$ y $V = \text{rot } U$ para algún campo vectorial U derivable con continuidad (en todo S).
   (II) Existe un campo escalar φ tal que $\nabla\varphi$ es derivable con continuidad y

$$V = \text{grad } \varphi \qquad \text{y} \qquad \nabla^2\varphi = 0 \quad \text{en todo } S.$$

   a) Demostrar que (I) implica (II). Dicho de otro modo, un campo vectorial que es a la vez irrotacional y solenoidal en S es el gradiente de una función armónica en S.
   b) Demostrar que (II) implica (I), o dar un contraejemplo.

8. Supongamos todos los campos vectoriales que se consideren derivables con continuidad. Sea $H = F + G$, donde F es solenoidal y G irrotacional. Existe entonces un campo vectorial U tal que $F = \text{rot } U$ y un campo escalar φ tal que $G = \nabla\varphi$ en S. Demostrar que U y φ satisfacen en S la siguiente ecuación diferencial en derivadas parciales:

$$\nabla^2\varphi = \text{div } H, \qquad \text{grad}(\text{div } U) - \nabla^2 U = \text{rot } H.$$

Observación. Este ejercicio es de gran utilidad, a causa de que puede demostrarse que todo campo vectorial H derivable con continuidad en S puede expresarse en la forma $H = F + G$ donde F es solenoidal y G irrotacional.

9. Sea $H(x, y, z) = x^2yi + y^2zj + z^2xk$. Hallar los campos vectoriales F y G, donde F es un rotacional y G un gradiente, de modo que se verifique $H = F + G$.

10. Sean u y v dos campos escalares derivables con continuidad en un intervalo abierto de E_3.
a) Demostrar que existe un campo vectorial F tal que $\nabla u \times \nabla v = \text{rot } F$ en todo R.
b) Determinar si cualquiera de los tres campos vectoriales siguientes pueden o no ser utilizados como F en la parte a): I) $\nabla(uv)$; II) $u\nabla v$; III) $v\nabla u$.
c) Si $u(x, y, z) = x^3 - y^3 + z^3$ y $v(x, y, z) = x + y + z$, calcular la integral de superficie

$$\iint_S \nabla u \times \nabla v \cdot n \, dS,$$

en donde S es la semiesfera $x^2 + y^2 + z^2 = 1$, $z \geq 0$, y n es la normal unitaria con un componente z no negativo.

12.18 Extensiones del teorema de Stokes

El teorema de Stokes puede extenderse a superficies regulares simples más generales. Si T es una región múltiplemente conexa parecida a la de la figura

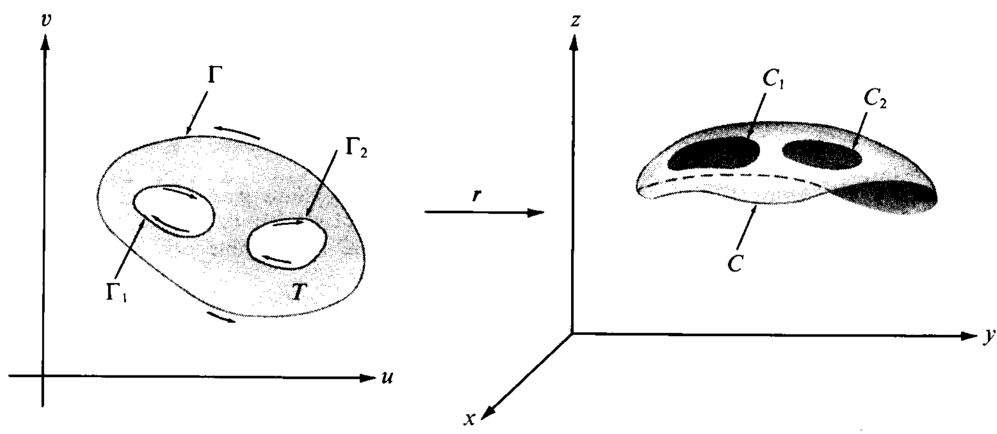


FIGURA 12.14 *Extensión del teorema de Stokes a superficies que son imágenes uno a uno de regiones múltiplemente conexas.*

12.14 (con un número finito de agujeros), la imagen uno a uno $S = \boldsymbol{r}(T)$ contendrá el mismo número de agujeros que T. Para extender el teorema de Stokes a tales superficiales seguiremos el mismo tipo de razonamiento que en la demostración precedente, excepto que usaremos el teorema de Green para regiones múltiplemente conexas (teorema 11.12). En lugar de la integral de línea que aparece en (12.27) precisamos una suma de integrales de línea, con signos adecuados, tomadas sobre las imágenes de las curvas que constituyen la frontera de T. Por ejemplo, si T tiene dos agujeros, como en la figura 12.14, y las curvas frontera Γ, Γ_1 y Γ_2 son recorridas en la dirección indicada, la identidad del teorema de Stokes toma la forma

$$\iint_S (\operatorname{rot} \boldsymbol{F}) \cdot \boldsymbol{n}\, dS = \oint_C \boldsymbol{F} \cdot d\boldsymbol{\rho} + \oint_{C_1} \boldsymbol{F} \cdot d\boldsymbol{\rho}_1 + \oint_{C_2} \boldsymbol{F} \cdot d\boldsymbol{\rho}_2,$$

donde C, C_1 y C_2 son las imágenes de Γ, Γ_1 y Γ_2 respectivamente, y $\boldsymbol{\rho}$, $\boldsymbol{\rho}_1$ y $\boldsymbol{\rho}_2$ son las funciones compuestas $\boldsymbol{\rho}(t) = \boldsymbol{r}[\boldsymbol{\gamma}(t)]$, $\boldsymbol{\rho}_1(t) = \boldsymbol{r}[\boldsymbol{\gamma}_1(t)]$, $\boldsymbol{\rho}_2(t) = \boldsymbol{r}[\boldsymbol{\gamma}_2(t)]$. Las funciones $\boldsymbol{\gamma}$, $\boldsymbol{\gamma}_1$ y $\boldsymbol{\gamma}_2$ son las que describen Γ, Γ_1 y Γ_2 en las direcciones indicadas. Las curvas C, C_1 y C_2 serán recorridas en las direcciones inducidas por la aplicación $\boldsymbol{r}$ a partir de Γ, Γ_1 y Γ_2.

El teorema de Stokes puede también extenderse a algunas superficies regulares no simples (pero no a todas). Veamos algunos ejemplos

Consideremos primero el cilindro dibujado en la figura 12.15. Es la reunión de dos superficies paramétricas regulares simples S_1 y S_2, imágenes de dos rectángulos adyacentes T_1 y T_2 a través de las aplicaciones $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$, respectivamente. Si $\boldsymbol{\gamma}_1$ describe la frontera Γ_1, positivamente orientada, de T_1 y $\boldsymbol{\gamma}_2$ la frontera Γ_2 de T_2 asimismo orientada positivamente, las funciones $\boldsymbol{\rho}_1$ y $\boldsymbol{\rho}_2$ definidas por

$$\boldsymbol{\rho}_1(t) = \boldsymbol{r}_1[\boldsymbol{\gamma}_1(t)], \qquad \boldsymbol{\rho}_2(t) = \boldsymbol{r}_2[\boldsymbol{\gamma}_2(t)]$$

describen las imágenes C_1 y C_2, de Γ_1 y Γ_2 respectivamente. En este ejemplo las representaciones $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$ pueden elegirse de modo que estén de acuerdo en la inter-

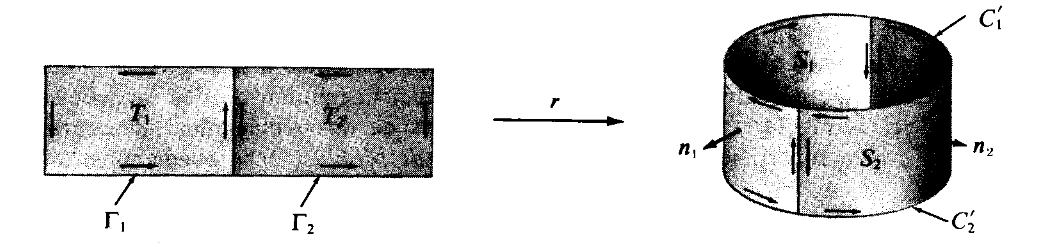


FIGURA 12.15 *Extensión del teorema de Stokes a un cilindro.*

sección $\Gamma_1 \cap \Gamma_2$. Si aplicamos el teorema de Stokes a cada parte S_1 y S_2 y sumamos las dos identidades, obtenemos:

$$(12.51)\quad \iint_{S_1} (\text{rot } \boldsymbol{F}) \cdot \boldsymbol{n}_1 \, dS + \iint_{S_2} (\text{rot } \boldsymbol{F}) \cdot \boldsymbol{n}_2 \, dS = \int_{C_1} \boldsymbol{F} \cdot d\boldsymbol{\rho}_1 + \int_{C_2} \boldsymbol{F} \cdot d\boldsymbol{\rho}_2 ,$$

donde $\boldsymbol{n}_1$ y $\boldsymbol{n}_2$ son las normales determinadas por los productos vectoriales de $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$, respectivamente.

Representemos con $\boldsymbol{r}$ la aplicación de $T_1 \cup T_2$ que coincide con $\boldsymbol{r}_1$ en T_1 y con $\boldsymbol{r}_2$ en T_2, y sea $\boldsymbol{n}$ la correspondiente normal unitaria determinada por el producto vectorial fundamental $\boldsymbol{r}$. Ya que las normales $\boldsymbol{n}_1$ y $\boldsymbol{n}_2$ coinciden en dirección en $S_1 \cap S_2$, $\boldsymbol{n}$ es igual a $\boldsymbol{n}_1$ en S_1 e igual a $\boldsymbol{n}_2$ en S_2. Por consiguiente, la suma de las integrales de superficie del primer miembro de (12.51) es igual a

$$\iint_{S_1 \cup S_2} (\text{rot } \boldsymbol{F}) \cdot \boldsymbol{n} \, dS .$$

En este ejemplo, las representaciones $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$ pueden elegirse de modo que $\boldsymbol{\rho}_1$ y $\boldsymbol{\rho}_2$ determinen direcciones *opuestas* en cada arco de la intersección $C_1 \cap C_2$, como indican las flechas en la figura 12.15. Las dos integrales de línea del segundo miembro de (12.51) pueden reemplazarse por una suma de integrales de línea a lo largo de las dos circunferencias C'_1 y C'_2 que forman el borde superior e inferior de $S_1 \cup S_2$, puesto que las integrales de línea a lo largo de cada arco de la intersección $C_1 \cap C_2$ se reducen. Por lo tanto, la ecuación (12.51) puede escribirse en la forma

$$(12.52)\qquad \iint_{S_1 \cup S_2} (\text{rot } \boldsymbol{F}) \cdot \boldsymbol{n} \, dS = \int_{C_1'} \boldsymbol{F} \cdot d\boldsymbol{\rho}_1 + \int_{C_2'} \boldsymbol{F} \cdot d\boldsymbol{\rho}_2 ,$$

donde las integrales de línea se calculan en las direcciones deducidas de las de Γ_1 y Γ_2. Las dos circunferencias C'_1 y C'_2 forman la frontera completa de $S_1 \cup S_2$. La ecuación (12.52) expresa la integral de superficie de (rot $\boldsymbol{F}$) $\cdot$ $\boldsymbol{n}$ sobre $S_1 \cup S_2$ como una integral de línea sobre la frontera completa de $S_1 \cup S_2$. Esa ecuación es la extensión del teorema de Stokes a un cilindro.

Supongamos ahora que aplicamos los mismos conceptos a la superficie dibujada en la figura 12.16. Esta superficie es también la reunión de dos superficies paramétricas regulares simples S_1 y S_2, imágenes de dos rectángulos adyacentes T_1 y T_2. Esta superficie especial se llama *banda de Möbius*(*); se puede construir

(*) De A. F. Möbius (1790-1868), discípulo de Gauss. A la edad de 26 años fue designado profesor de Astronomía en Leipzig, cargo que desempeñó hasta su muerte. Trabajó en Mecánica celeste, pero sus investigaciones más importantes versaron sobre temas de Geometría y de Teoría de números.

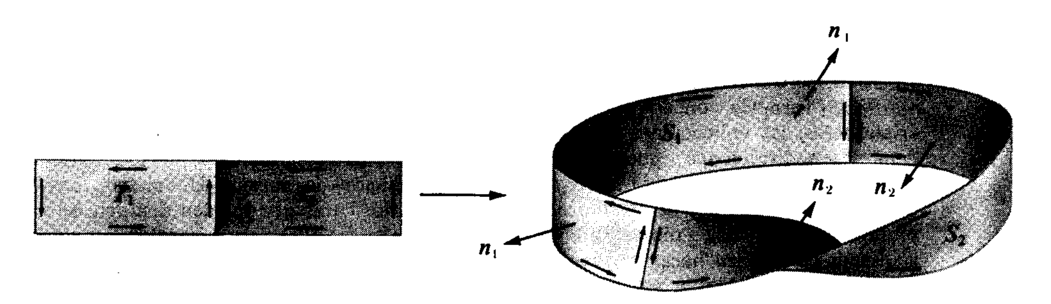


FIGURA 12.16 *Banda de Möbius considerada como la reunión de dos superficies paramétricas simples. El teorema de Stokes no se extiende a una banda de Möbius.*

fácilmente un modelo con una tira rectangular de papel dando a uno de los extremos una media vuelta y soldando los dos bordes. Definimos ρ_1, ρ_2, C_1 y C_2 para la banda de Möbius como antes se hizo para el cilindro. El borde de $S_1 \cup S_2$ en este caso es una curva cerrada simple C', y no dos. Esta curva es el borde completo de la banda de Möbius.

Si aplicamos el teorema de Stokes a cada una de las partes S_1 y S_2 como se hizo para el cilindro, obtenemos la ecuación (12.51). Pero si intentamos mantener las dos integrales de superficie y las dos integrales de línea como antes, nos encontramos con dos dificultades. Primera, las dos normales $\boldsymbol{n}_1$ y $\boldsymbol{n}_2$ no coinciden en dirección en toda la intersección $C_1 \cap C_2$. (Véase figura 12.16.) Por consiguiente, no podemos definir una normal $\boldsymbol{n}$ para toda la superficie tomando $\boldsymbol{n} = \boldsymbol{n}_1$ en S_1 y $\boldsymbol{n} = \boldsymbol{n}_2$ en S_2, como hicimos en el cilindro. Esto no es grave inconveniente, no obstante, ya que podemos definir $\boldsymbol{n} = \boldsymbol{n}_1$ en S_1 y en $C_1 \cap C_2$, y definir entonces $\boldsymbol{n} = \boldsymbol{n}_2$ en el resto. Esto nos da una normal discontinua, pero las discontinuidades así introducidas constituyen un conjunto de medida nula en el plano uv y no afectan la existencia o el valor de la integral de superficie

$$\iint_{S_1 \cup S_2} (\text{rot } \boldsymbol{F}) \cdot \boldsymbol{n} \, dS .$$

Una dificultad más seria encontramos al intentar mantener las integrales de línea. En este ejemplo no es posible elegir las aplicaciones $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$ de manera que ρ_1 y ρ_2 determinen direcciones opuestas en cada uno de los arcos de la intersección $C_1 \cap C_2$. Esto se indica con flechas en la figura 12.16, uno de esos arcos es recorrido dos veces en la misma dirección. Sobre este arco las correspondientes integrales de línea no se reducirán necesariamente como ocurrió en el cilindro. Por consiguiente, la suma de las integrales de línea de (12.51) no son necesariamente iguales a la integral de línea sobre la frontera completa de $S_1 \cup S_2$, y el teorema de Stokes no puede extenderse a la banda de Möbius.

Observación. El cilindro y la banda de Möbius son ejemplos de superficies *orientables* y *no orientables*, respectivamente. No intentaremos definir estos conceptos con precisión, pero mencionaremos algunas de sus diferencias. Para una superficie orientable $S_1 \cup S_2$ formada con dos superficies paramétricas regulares simples como las descritas antes, las aplicaciones $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$ pueden siempre elegirse de modo que $\boldsymbol{\rho}_1$ y $\boldsymbol{\rho}_2$ determinen direcciones opuestas sobre cada arco de la intersección $C_1 \cap C_2$. Para una superficie no orientable no es posible tal definición de normal. Un modelo en papel de una superficie orientable siempre presenta dos caras que pueden distinguirse pintándolas con dos colores diferentes. Las superficies no orientables tienen tan sólo una cara. Para hacer un estudio riguroso de éstas y otras propiedades de las superficies orientables y no orientables, puede consultarse cualquier libro de Topología combinatoria. El teorema de Stokes puede extenderse a las superficies orientables mediante un procedimiento parecido al seguido antes para el cilindro.

Otra superficie orientable es la esfera dibujada en la figura 12.17. Es la reunión de dos superficies paramétricas simples (hemisferios) S_1 y S_2, que pueden considerarse imágenes de un disco circular del plano xy a través de las aplicaciones $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$, respectivamente. Damos a $\boldsymbol{r}$, $\boldsymbol{\rho}_1$, $\boldsymbol{\rho}_2$, C_1, C_2 el mismo significado que en los ejemplos anteriores. En este caso las curvas C_1 y C_2 están identificadas por la aplicación $\boldsymbol{r}$ (coinciden a lo largo del ecuador), y la superficie $S_1 \cup S_2$ se llama *cerrada*. Además, $\boldsymbol{n}_1$ y $\boldsymbol{n}_2$ pueden elegirse de modo que las direcciones determinadas por $\boldsymbol{\rho}_1$ y $\boldsymbol{\rho}_2$ sean opuestas en C_1 y C_2, como se indica con flechas en la figura 12.17. (Esto ocurre porque $S_1 \cup S_2$ es orientable.) Si aplicamos el

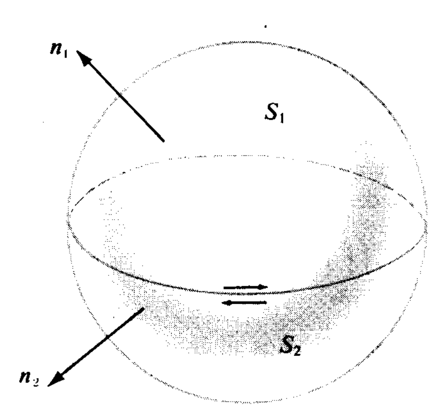


FIGURA 12.17 *Extensión del teorema de Stokes a una esfera.*

teorema de Stokes a cada hemisferio y sumamos los resultados obtenemos la ecuación (12.51), como antes. Las normales $\boldsymbol{n}_1$ y $\boldsymbol{n}_2$ coinciden en la intersección $C_1 \cap C_2$, y podemos reunir las integrales sobre S_1 y S_2 en una sobre toda la esfera.

Las dos integrales de línea del segundo miembro de (12.51) se reducen, y nos queda la fórmula

$$\iint_{S_1 \cup S_2} (\text{ rot } \boldsymbol{F}) \cdot \boldsymbol{n} \, dS = 0 .$$

Esto es válido no tan sólo para una esfera, sino para toda superficie orientable cerrada.

12.19 Teorema de la divergencia (teorema de Gauss)

El teorema de Stokes expresa una relación entre una integral extendida a una superficie y una integral de línea tomada sobre la curva o curvas que constituyen la frontera de tal superficie. El teorema de la divergencia expresa una relación entre una integral triple extendida a un sólido y una integral de superficie tomada sobre la frontera de ese sólido.

TEOREMA 12.6. TEOREMA DE LA DIVERGENCIA. *Si V es un sólido en E_3 limitado por una superficie orientable S, si $\boldsymbol{n}$ es la normal unitaria exterior a S y si $\boldsymbol{F}$ es un campo vectorial definido en V, entonces tenemos*

$$\iiint_V (\text{div } \boldsymbol{F}) \, dx \, dy \, dz = \iint_S \boldsymbol{F} \cdot \boldsymbol{n} \, dS . \tag{12.53}$$

Observación. Si expresamos $\boldsymbol{F}$ y $\boldsymbol{n}$ en función de sus componentes

$$\boldsymbol{F}(x, y, z) = P(x, y, z)\boldsymbol{i} + Q(x, y, z)\boldsymbol{j} + R(x, y, z)\boldsymbol{k}$$

y

$$\boldsymbol{n} = \cos \alpha \, \boldsymbol{i} + \cos \beta \, \boldsymbol{j} + \cos \gamma \, \boldsymbol{k} ,$$

la ecuación (12.53) puede entonces ponerse en la forma:

$$\iiint_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx \, dy \, dz = \iint_S (P \cos \alpha + Q \cos \beta + R \cos \gamma) \, dS . \tag{12.54}$$

Demostración. Bastará establecer las tres ecuaciones

$$\iiint_V \frac{\partial P}{\partial x} \, dx \, dy \, dz = \iint_S P \cos \alpha \, dS ,$$

$$\iiint_V \frac{\partial Q}{\partial y}\, dx\, dy\, dz = \iint_S Q \cos \beta\, dS,$$

$$\iiint_V \frac{\partial R}{\partial z}\, dx\, dy\, dz = \iint_S R \cos \gamma\, dS,$$

y sumar los resultados para obtener (12.54). Comenzamos por la tercera de esas fórmulas y la demostramos para sólidos de tipo ciertamente especial.

Supongamos que V es un conjunto de puntos (x, y, z) que satisfacen una relación de la forma

$$g(x, y) \leq z \leq f(x, y) \qquad \text{para} \quad (x, y) \text{ en } T,$$

siendo T una región conexa del plano xy, y f y g funciones continuas en T, con la condición $g(x, y) \leq f(x, y)$ para cada punto (x, y) en T. Geométricamente, esto significa que T es la proyección de V en el plano xy. Toda recta paralela al eje z que atraviese T corta al sólido V a lo largo de un segmento rectilíneo que une la superficie $z = g(x, y)$ a la $z = f(x, y)$. La superficie frontera S consta de un casquete superior S_1, dado en la forma explícita $z = f(x, y)$; otro inferior S_2 dado por $z = g(x, y)$; y en algunos casos por una porción de cilindro S_3 engendrado por una recta que se mueve a lo largo de la frontera de T manteniéndose paralela al eje z. La normal exterior a S tiene componente z no negativa en S_1 y no positiva en S_2 y es paralela al plano xy en S_3. Los sólidos de este tipo se llaman «proyectables-xy». (En la figura 12.18 se muestra un ejemplo.) En él se incluyen todos

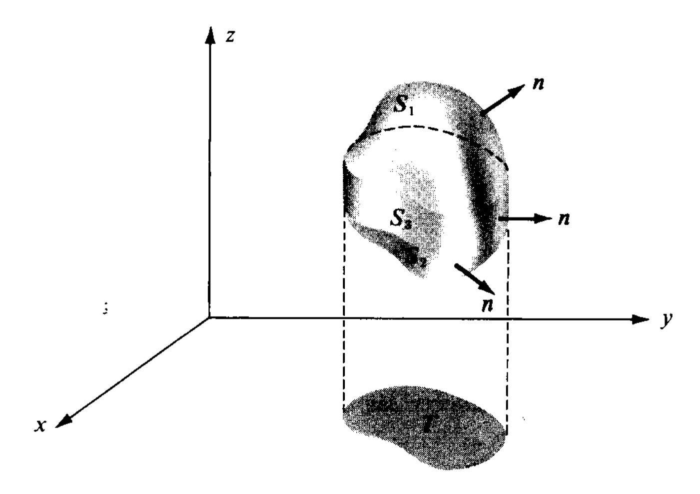


FIGURA 12.18 *Ejemplo de sólido proyectable-xy.*

los sólidos convexos (por ejemplo, esferas, elipsoides, cubos) y otros muchos que no son convexos (por ejemplo, el toro con eje paralelo al z).

La idea de la demostración es sencillísima. Expresamos la integral triple como una doble extendida a la proyección T. Entonces demostramos que esta integral doble tiene el mismo valor que la integral de superficie citada en el enunciado. Comencemos con la fórmula

$$\iiint_V \frac{\partial R}{\partial z}\, dx\, dy\, dz = \iint_T \left[\int_{g(x,y)}^{f(x,y)} \frac{\partial R}{\partial z}\, dz\right] dx\, dy\,.$$

La integral unidimensional respecto a z puede calcularse mediante el segundo teorema fundamental del cálculo, dándonos

$$(12.55)\qquad \iiint_V \frac{\partial R}{\partial z}\, dx\, dy\, dz = \iint_T \{R[x, y, f(x, y)] - R[x, y, g(x, y)]\}\, dx\, dy\,.$$

Para la integral de superficie podemos escribir:

$$(12.56)\quad \iint_S R \cos\gamma\, dS = \iint_{S_1} R \cos\gamma\, dS + \iint_{S_2} R \cos\gamma\, dS + \iint_{S_3} R \cos\gamma\, dS\,.$$

Sobre S_3 la normal $\boldsymbol{n}$ es paralela al plano xy, de modo que $\cos\gamma = 0$, y la integral sobre S_3 es nula. Sobre la superficie S_1 usamos la representación

$$\boldsymbol{r}(x, y) = x\boldsymbol{i} + y\boldsymbol{j} + f(x, y)\boldsymbol{k}\,,$$

y sobre S_2

$$\boldsymbol{r}(x, y) = x\boldsymbol{i} + y\boldsymbol{j} + g(x, y)\boldsymbol{k}\,.$$

En S_1 la normal $\boldsymbol{n}$ tiene la misma dirección que el producto vectorial fundamental $\partial\boldsymbol{r}/\partial x \times \partial\boldsymbol{r}/\partial y$, así podemos escribir [véase ecuación (12.25)]:

$$\iint_{S_1} R \cos\gamma\, dS = \iint_{S_1} R\, dx \wedge dy = \iint_T R[x, y, f(x, y)]\, dx\, dy\,.$$

En S_2 la normal $\boldsymbol{n}$ tiene dirección opuesta a la de $\partial\boldsymbol{r}/\partial x \times \partial\boldsymbol{r}/\partial y$ de modo que, en virtud de (12.26), tenemos,

$$\iint_{S_2} R \cos\gamma\, dS = -\iint_{S_2} R\, dx \wedge dy = -\iint_T R[x, y, g(x, y)]\, dx\, dy\,.$$

Por consiguiente, la ecuación (12.56) se convierte en

$$\iint_S R \cos \gamma \, dS = \iint_T \{R[x, y, f(x, y)] - R[x, y, g(x, y)]\} \, dx \, dy .$$

Comparando ésta con la (12.55) vemos que

$$\iiint_V \frac{\partial R}{\partial z} \, dx \, dy \, dz = \iint_S R \cos \gamma \, dS .$$

En la demostración anterior la hipótesis de que V es proyectable-xy nos permite expresar la integral triple extendida a V como una integral doble sobre su proyección T sobre el plano xy. Es evidente que si V es proyectable-yz podemos razonar del mismo modo y demostrar la identidad

$$\iiint_V \frac{\partial P}{\partial x} \, dx \, dy \, dz = \iint_S P \cos \alpha \, dS ;$$

y si V es proyectable-xz obtenemos:

$$\iiint_V \frac{\partial Q}{\partial y} \, dx \, dy \, dz = \iint_S Q \cos \beta \, dS .$$

Así vemos que el teorema de la divergencia es válido para todos los sólidos proyectables sobre los tres planos coordenados; en particular, para todo sólido convexo.

Un toro con el eje paralelo al eje z es proyectable-xy pero no proyectable-xz ni proyectable-yz. Para extender el teorema de la divergencia a ese sólido cortamos el toro en cuatro partes iguales por medio de planos que pasan por su eje y paralelos a los planos xz e yz, respectivamente, y aplicamos el teorema a cada una de las partes. La integral triple sobre el toro completo es la suma de las integrales triples sobre las cuatro partes. Cuando sumamos las integrales de superficie sobre las cuatro partes encontramos que las aportaciones de las caras comunes a las partes adyacentes se reducen unas con otras, puesto que las normales exteriores tienen direcciones opuestas sobre un par de tales caras. Por lo tanto, la suma de las integrales de superficie sobre las cuatro partes es igual a la integral de superficie sobre el toro completo. Este ejemplo hace ver cómo el teorema de la divergencia puede extenderse a ciertos sólidos no convexos.

12.20 Aplicaciones del teorema de la divergencia

Los conceptos de rotacional y divergencia de un campo vectorial $\boldsymbol{F} = P\boldsymbol{i} + Q\boldsymbol{j} + R\boldsymbol{k}$ se introdujeron en la sección 12.12 mediante las fórmulas

$$\text{div}\,\boldsymbol{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \tag{12.57}$$

y

$$\text{rot}\ \boldsymbol{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\boldsymbol{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\boldsymbol{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\boldsymbol{k}. \tag{12.58}$$

Para calcular div $\boldsymbol{F}$ y rot $\boldsymbol{F}$ a partir de esas fórmulas se necesita el conocimiento de los componentes de $\boldsymbol{F}$. Esos componentes, a su vez, dependen de la elección de los ejes coordenados en E_3. Un cambio en la posición de los ejes coordenados significaría un cambio en los componentes de $\boldsymbol{F}$ y, como cabe esperar, el correspondiente cambio en las funciones div $\boldsymbol{F}$ y rot $\boldsymbol{F}$. Con la ayuda del teorema de Stokes y del de la divergencia podemos obtener fórmulas para la divergencia y el rotacional en las que no intervienen los componentes de $\boldsymbol{F}$. Estas fórmulas demuestran que el rotacional y la divergencia representan propiedades intrínsecas del campo vectorial $\boldsymbol{F}$ y no dependen de la elección que se haga de los ejes coordenados. Vamos a tratar primero la fórmula para la divergencia.

TEOREMA 12.7. *Sea $V(t)$ una esfera de radio $t > 0$ con centro en el punto $\boldsymbol{a}$ de E_3, y representemos con $S(t)$ la frontera de $V(t)$. Sea $\boldsymbol{F}$ un campo vectorial derivable con continuidad en $V(t)$. Entonces si $|V(t)|$ representa el volumen de $V(t)$, y $\boldsymbol{n}$ la normal unitaria exterior a S, tenemos*

$$\text{div}\,\boldsymbol{F}(\boldsymbol{a}) = \lim_{t\to 0} \frac{1}{|V(t)|} \iint_{S(t)} \boldsymbol{F}\cdot\boldsymbol{n}\,dS. \tag{12.59}$$

Demostración. Sea $\varphi = \text{div}\,\boldsymbol{F}$. Si $\epsilon > 0$ tenemos que encontrar un $\delta > 0$ tal que

$$\left|\varphi(\boldsymbol{a}) - \frac{1}{|V(t)|}\iint_{S(t)} \boldsymbol{F}\cdot\boldsymbol{n}\,dS\right| < \epsilon \quad \text{siempre que} \quad 0 < t < \delta. \tag{12.60}$$

Puesto que φ es continua en $\boldsymbol{a}$, para el ϵ dado existe una 3-esfera $B(\boldsymbol{a}; h)$ tal que

$$|\varphi(\boldsymbol{x}) - \varphi(\boldsymbol{a})| < \frac{\epsilon}{2} \quad \text{siempre que} \quad \boldsymbol{x} \in B(\boldsymbol{a}; h).$$

Por consiguiente, si escribimos $\varphi(\boldsymbol{a}) = \varphi(\boldsymbol{x}) + [\varphi(\boldsymbol{a}) - \varphi(\boldsymbol{x})]$ e integramos ambos miembros de esta ecuación sobre la esfera $V(t)$ de radio $t < h$, encontramos

$$\varphi(\boldsymbol{a})\,|V(t)| = \iiint_{V(t)} \varphi(\boldsymbol{x})\,dx\,dy\,dz + \iiint_{V(t)} [\varphi(\boldsymbol{a}) - \varphi(x)]\,dx\,dy\,dz\,.$$

Si aplicamos el teorema de la divergencia a la primera integral triple del segundo miembro y pasamos este término al primer miembro, obtenemos la relación,

$$\left|\varphi(\boldsymbol{a})\,|V(t)| - \iint_{S(t)} \boldsymbol{F}\cdot\boldsymbol{n}\,dS\right| \leq \iiint_{V(t)} |\varphi(\boldsymbol{a}) - \varphi(\boldsymbol{x})|\,dx\,dy\,dz \leq \frac{\epsilon}{2}\,|V(t)| < \epsilon\,|V(t)|\,.$$

Cuando dividimos esta desigualdad por $|V(t)|$ vemos que (12.60) es cierta para $\delta = h$. Esto demuestra el teorema.

En la demostración anterior no hacemos uso especial del hecho de que $V(t)$ fuese una esfera. El mismo teorema subsiste si, en lugar de esferas, utilizamos cualquier conjunto de sólidos $V(t)$ para los que el teorema de la divergencia es válido, con tal que esos sólidos contengan el punto $\boldsymbol{a}$ y tiendan hacia $\boldsymbol{a}$ cuando $t \to 0$. Por ejemplo, cada $V(t)$ podría ser un cubo inscrito en una esfera de radio t en torno de $\boldsymbol{a}$; se aplicaría exactamente la misma demostración.

El teorema 12.7 puede también justificarse con razones físicas. Supongamos que $\boldsymbol{F}$ representa el vector densidad de flujo de una corriente estacionaria. La integral de superficie $\iint_{S(t)} \boldsymbol{F}\cdot\boldsymbol{n}\,dS$ mide entonces la masa total de fluido que atraviesa S en la unidad de tiempo y en la dirección de $\boldsymbol{n}$. El cociente $\iint_{S(t)} \boldsymbol{F}\cdot\boldsymbol{n}\,dS/|V(t)|$ representa la masa por unidad de volumen que fluye a través de S en la unidad de tiempo en la dirección de $\boldsymbol{n}$. Cuando $t \to 0$, el límite de ese cociente es la divergencia de $\boldsymbol{F}$ en $\boldsymbol{a}$. Luego la divergencia en $\boldsymbol{a}$ puede interpretarse como el coeficiente de variación de masa por unidad de volumen y por unidad de tiempo en $\boldsymbol{a}$.

En algunos libros de Análisis vectorial, la ecuación (12.59) se toma como *definición* de divergencia. Esto permite inmediatamente asignar un significado físico a la divergencia. Además, en la fórmula (12.59) no intervienen los componentes de $\boldsymbol{F}$. Por lo tanto, es válida en cualquier sistema de coordenadas. Si elegimos como $V(t)$ un cubo con sus aristas paralelas a los ejes xyz y con centro en $\boldsymbol{a}$, podemos utilizar la ecuación (12.59) para deducir la fórmula (12.57) que expresa la div $\boldsymbol{F}$ en función de los componentes de $\boldsymbol{F}$. Este procedimiento se esboza en el ejercicio 14 de la sección 12.21.

Existe una fórmula análoga a la (12.59) que a veces se usa como otra definición de rotacional. Dicha fórmula es:

$$\text{(12.61)} \qquad \operatorname{rot} \boldsymbol{F}(\boldsymbol{a}) = \lim_{t\to 0} \frac{1}{|V(t)|} \iint_{S(t)} \boldsymbol{n} \times \boldsymbol{F}\, dS,$$

donde $V(t)$ y $S(t)$ tienen el mismo significado que en el teorema 12.7. La integral de superficie que aparece en el segundo miembro tiene un integrando vectorial. Tales integrales pueden definirse en función de los componentes. La demostración de (12.61) es análoga a la del teorema 12.7.

Existe otra fórmula en la que interviene el rotacional y que puede deducirse de (12.61) o bien directamente. Tal forma establece que

$$\text{(12.62)} \qquad \boldsymbol{n} \cdot \operatorname{rot} \boldsymbol{F}(\boldsymbol{a}) = \lim_{t\to 0} \frac{1}{|S(t)|} \oint_{C(t)} \boldsymbol{F} \cdot d\boldsymbol{\alpha}.$$

En esta fórmula, $S(t)$ es un disco circular de radio t y centro en $\boldsymbol{a}$, y $|S(t)|$ representa su área. El vector $\boldsymbol{n}$ es la normal unitaria a $S(t)$, y $\boldsymbol{\alpha}$ es la función que describe $C(t)$ en una dirección tal que vista desde el extremo de $\boldsymbol{n}$ parece contraria a la de las agujas del reloj. El campo vectorial $\boldsymbol{F}$ se supone derivable con continuidad en $S(t)$. Puede darse una demostración de (12.62) con el mismo método usado para demostrar (12.59). Ponemos $\varphi(\boldsymbol{x}) = \boldsymbol{n} \cdot \operatorname{rot} \boldsymbol{F}(\boldsymbol{x})$ y razonamos como antes, salvo que empleamos integrales de superficie en lugar de integrales triples y el teorema de Stokes en lugar del de la divergencia.

Si $\boldsymbol{F}$ es un campo de velocidad, la integral de línea sobre $C(t)$ se llama circulación de $\boldsymbol{F}$ a lo largo de $C(t)$; el límite de (12.62) representa la circulación por unidad de área en el punto $\boldsymbol{a}$. De este modo, $\boldsymbol{n} \cdot \operatorname{rot} \boldsymbol{F}(\boldsymbol{a})$ puede considerarse como una «densidad de circulación» de $\boldsymbol{F}$ en el punto $\boldsymbol{a}$, con respecto a un plano perpendicular a $\boldsymbol{n}$ en $\boldsymbol{a}$.

Cuando $\boldsymbol{n}$ toma sucesivamente los valores $\boldsymbol{i}$, $\boldsymbol{j}$, y $\boldsymbol{k}$, los productos interiores $\boldsymbol{i} \cdot \operatorname{rot} \boldsymbol{F}$, $\boldsymbol{j} \cdot \operatorname{rot} \boldsymbol{F}$, y $\boldsymbol{k} \cdot \operatorname{rot} \boldsymbol{F}$ son los componentes de rot $\boldsymbol{F}$ en coordenadas rectangulares. Cuando la ecuación (12.61) se toma como punto de partida de la definición de rotacional, la fórmula (12.58) de los componentes rectangulares de rot $\boldsymbol{F}$ pueden deducirse de (12.62) exactamente del mismo modo.

12.21 Ejercicios

1. Sea S la superficie del cubo unidad $0 \le x \le 1$, $0 \le y \le 1$, $0 \le z \le 1$, y sea $\boldsymbol{n}$ la normal unitaria exterior a S. Si $\boldsymbol{F}(x, y, z) = x^2\boldsymbol{i} + y^2\boldsymbol{j} + z^2\boldsymbol{k}$, empleamos el teorema de la divergencia para calcular la integral de superficie $\iint_S \boldsymbol{F} \cdot \boldsymbol{n}\, dS$. Comprobar el resultado calculando la integral de superficie directamente.
2. Se corta la esfera $x^2 + y^2 + z^2 = 25$ por el plano $z = 3$. La parte menor es un sólido V limitado por una superficie S_0 constituida por dos partes, una esférica S_1 y otra

plana S_2. Si la normal unitaria exterior a V es $\cos\alpha\, \boldsymbol{i} + \cos\beta\, \boldsymbol{j} + \cos\gamma\, \boldsymbol{k}$, calcular el valor de la integral de superficie

$$\iint_S (xz\cos\alpha + yz\cos\beta + \cos\gamma)\, dS$$

si a) S es el casquete esférico S_1, b) S es la base plana S_2, c) S es la frontera completa S_0. Resolver la parte *c*) con los resultados de las partes a) y b), y también usando el teorema de la divergencia.

3. Sea $\boldsymbol{n} = \cos\alpha\, \boldsymbol{i} + \cos\beta\, \boldsymbol{j} + \cos\gamma\, \boldsymbol{k}$ la normal unitaria exterior a una superficie cerrada S que limita un sólido homogéneo V del tipo descrito en el teorema de la divergencia. Supongamos que el centro de gravedad $(\bar{x}, \bar{y}, \bar{z})$ y el volumen $|V|$ de V son conocidos. Calcular las siguientes integrales de superficie en función de $|V|$ y de $\bar{x}, \bar{y}, \bar{z}$.

a) $\iint_S (x\cos\alpha + y\cos\beta + z\cos\gamma)\, dS$.

b) $\iint_S (xz\cos\alpha + 2yz\cos\beta + 3z^2\cos\gamma)\, dS$.

c) $\iint_S (y^2\cos\alpha + 2xy\cos\beta - xz\cos\gamma)\, dS$.

d) Expresar $\iint_S (x^2 + y^2)(x\boldsymbol{i} + y\boldsymbol{j}) \cdot \boldsymbol{n}\, dS$ en función del volumen $|V|$ y un momento de inercia del sólido.

En los ejercicios del 4 al 10, $\partial f/\partial n$ y $\partial g/\partial n$ indican las derivadas direccionales de los campos escalares f y g en la dirección de la normal unitaria exterior $\boldsymbol{n}$ a una superficie cerrada S que limita un sólido V del tipo descrito en el teorema de la divergencia. Esto es, $\partial f/\partial n = \nabla f \cdot \boldsymbol{n}$ y $\partial g/\partial n = \nabla g \cdot \boldsymbol{n}$. En cada uno de los ejercicios demostrar la igualdad que se da. Se puede suponer la continuidad de todas las derivadas que intervienen.

4. $$\iint_S \frac{\partial f}{\partial n}\, dS = \iiint_V \nabla^2 f\, dx\, dy\, dz.$$

5. $$\iint_S \frac{\partial f}{\partial n}\, dS = 0 \quad \text{siempre que } f \text{ sea armónica en } V.$$

6. $$\iint_S f\frac{\partial g}{\partial n}\, dS = \iiint_V f\nabla^2 g\, dx\, dy\, dz + \iiint_V \nabla f \cdot \nabla g\, dx\, dy\, dz.$$

7. $$\iint_S \left(f\frac{\partial g}{\partial n} - g\frac{\partial f}{\partial n}\right) dS = \iiint_V (f\nabla^2 g - g\nabla^2 f)\, dx\, dy\, dz.$$

8. $$\iint_S f\frac{\partial g}{\partial n}\, dS = \iint_S g\frac{\partial f}{\partial n}\, dS \quad \text{si } f \text{ y } g \text{ son ambas armónicas en V.}$$

9. $$\iint_S f\frac{\partial f}{\partial n}\,dS = \iiint_V |\nabla f|^2\,dx\,dy\,dz \qquad \text{si } f \text{ es armónica en } V.$$

10. $$\nabla^2 f(\boldsymbol{a}) = \lim_{t\to 0}\frac{1}{|V(t)|}\iint_{S(t)}\frac{\partial f}{\partial n}\,dS,$$ donde $V(t)$ es una esfera de radio r y centro en $\boldsymbol{a}$, $S(t)$ es la superficie de $V(t)$, y $|V(t)|$ es el volumen de $V(t)$.

11. Sean V una región convexa de E_3 cuya frontera es una superficie cerrada S y $\boldsymbol{n}$ la normal unitaria exterior a S. Sean $\boldsymbol{F}$ y $\boldsymbol{G}$ dos campos vectoriales derivables con continuidad tales que

$$\text{rot } \boldsymbol{F} = \text{rot } \boldsymbol{G} \qquad \text{y} \qquad \text{div } \boldsymbol{F} = \text{div } \boldsymbol{G} \qquad \text{en todo } V,$$

y que satisfacen

$$\boldsymbol{G}\cdot\boldsymbol{n} = \boldsymbol{F}\cdot\boldsymbol{n} \qquad \text{en toda la superficie } S.$$

Demostrar que $\boldsymbol{F} = \boldsymbol{G}$ en todo V. [*Indicación.* Sea $\boldsymbol{H} = \boldsymbol{F} - \boldsymbol{G}$, encontrar un campo escalar f tal que $\boldsymbol{H} = \nabla f$, y usar una identidad adecuada para demostrar que $\iiint_V \|\nabla f\|^2\,dx\,dy\,dz = 0$. De esto se deduce que $\boldsymbol{H} = \boldsymbol{O}$ en V.]

12. Dados un campo vectorial $\boldsymbol{G}$ y dos campos escalares f y g, ambos derivables con continuidad en un sólido convexo V limitado por una superficie cerrada S. La normal unitaria exterior a S la representamos con $\boldsymbol{n}$. Demostrar que existe por lo menos un campo vectorial $\boldsymbol{F}$ que satisface las tres condiciones siguientes:

$$\text{rot } \boldsymbol{F} = \boldsymbol{G} \qquad \text{y} \qquad \text{div } \boldsymbol{F} = g \qquad \text{en } V, \qquad \boldsymbol{F}\cdot\boldsymbol{n} = f \quad \text{sobre } S.$$

13. Sea S una superficie paramétrica regular con la propiedad de que cada recta que pasa por P corta a S una vez a lo sumo. Representemos con $\Omega(S)$ el conjunto de rectas que pasando por P atraviesan S. (Véase figura 12.19.) El conjunto $\Omega(S)$ se llama *ángulo sólido* de vértice en P subtendido por S. Sea $\Sigma(a)$ la intersección de $\Omega(S)$ en la superficie de la esfera de radio a y centro en P. El cociente

$$\frac{\text{area de } \Sigma(a)}{a^2}$$

que se representa por $|\Omega(S)|$ y se utiliza como medida del ángulo sólido $\Omega(S)$.

a) Demostrar que este cociente es igual a la integral de superficie

$$\iint_S \frac{\boldsymbol{r}\cdot\boldsymbol{n}}{r^3}\,dS,$$

donde $\boldsymbol{r}$ es el radio vector que une P a un punto cualquiera de S, y $r = \|\boldsymbol{r}\|$. El vector $\boldsymbol{n}$ es la normal unitaria a S en la dirección de alejamiento de P. Esto muestra que el cociente $|\Omega(S)|$ es independiente del radio a. Por consiguiente, el ángulo sólido puede ser medido por el área de la intersección de $\Omega(S)$ y la esfera unidad en torno a P. [*Indicación.* Aplicar el teorema de la divergencia a la porción de $\Omega(S)$ comprendida entre S y $\Sigma(a)$.]

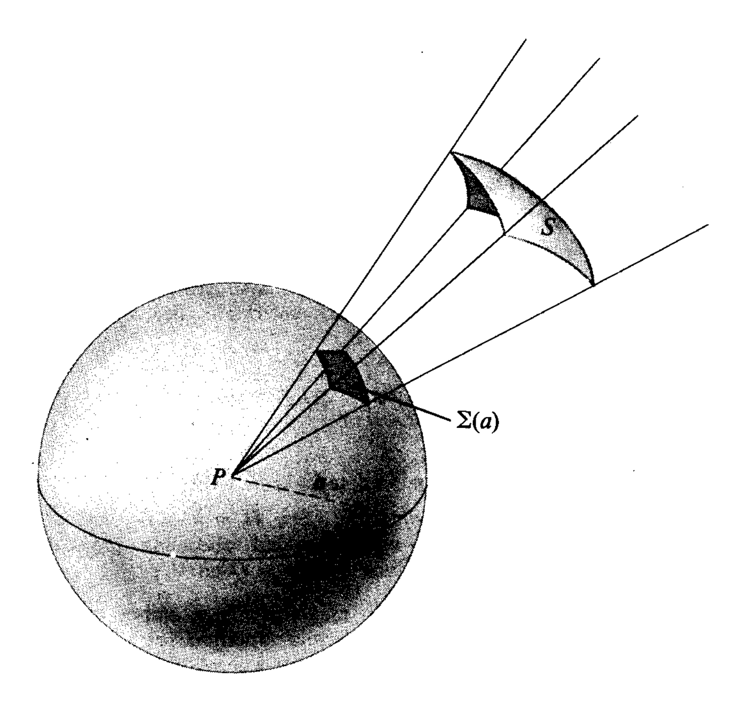


FIGURA 12.19 *Ángulo sólido* $\Omega(S)$ *con vértice en* P *subtendido por una superficie* S. *Su medida viene dada por el cociente* $|\Omega(S)| = \dfrac{\text{area de } \Sigma(a)}{a^2}$.

b) Dos planos se cortan a lo largo de un diámetro de una esfera con centro en P. El ángulo de intersección es θ, siendo $0 < \theta < \pi$. Sea S la parte menor de la superficie de la esfera interceptada por los dos planos. Demostrar que $|\Omega(S)| = 2\theta$.

14. Sean $V(t)$ un cubo de arista $2t$ y centro en $\boldsymbol{a}$, y $S(t)$ la frontera del citado cubo. Como de costumbre $\boldsymbol{n}$ es la normal unitaria exterior a $S(t)$ y sea $|V(t)|$ el volumen del cubo. Para un campo vectorial dado $\boldsymbol{F}$ diferenciable con continuidad en $\boldsymbol{a}$, suponer que existe el siguiente límite:

$$\lim_{t\to 0} \frac{1}{|V(t)|} \iint_{S(t)} \boldsymbol{F}\cdot\boldsymbol{n}\, dS,$$

y emplearlo como definición de div $\boldsymbol{F}(\boldsymbol{a})$. Elegidos unos ejes coordenados xyz paralelos a las aristas de $V(t)$, sean P, Q y R los componentes de $\boldsymbol{F}$ relativos a ese sistema de coordenadas. Demostrar que div $\boldsymbol{F}(\boldsymbol{a}) = D_1P(\boldsymbol{a}) + D_2Q(\boldsymbol{a}) + D_3R(\boldsymbol{a})$. [*Indicación.* Expresar la integral de superficie como suma de seis integrales dobles tomadas sobre las caras del cubo. Demostrar entonces que $1/|V(t)|$ multiplicado por la suma de las dos integrales dobles sobre las caras perpendiculares al eje z tiende al límite $D_3R(\boldsymbol{a})$ cuando $t \to 0$. Razonar de modo parecido para los restantes términos.]

15. Un campo escalar φ tiene la propiedad

$$\|\nabla\varphi\|^2 = 4\varphi \qquad \text{y} \qquad \operatorname{div}(\varphi\nabla\varphi) = 10\varphi .$$

Calcular la integral de superficie

$$\iint_S \frac{\partial \varphi}{\partial n}\, dS,$$

donde S es la superficie de la esfera unidad con centro en el origen, y $\partial\varphi/\partial n$ es la derivada direccional de φ en la dirección de la normal unitaria exterior a S.

PARTE 3

Temas especiales

13

FUNCIONES DE CONJUNTO Y PROBABILIDAD ELEMENTAL

13.1 Introducción histórica

Una disputa entre jugadores en 1654 llevó a dos famosos matemáticos franceses, Blaise Pascal y Pierre de Fermat, a la creación del Cálculo de probabilidades. Antoine Gombaud, caballero de Méré, noble francés interesado en cuestiones de juegos y apuestas, llamó la atención a Pascal respecto a una aparente contradicción en un popular juego de dados. El juego consistía en lanzar 24 veces un par de dados; y el problema en decidir si era lo mismo apostar la misma cantidad a favor o en contra de la aparición por lo menos de un «doble seis» en las 24 tiradas. Una regla del juego aparentemente bien establecida condujo a de Méré a creer que apostar por un doble seis en 24 tiradas era ventajoso, pero sus propios cálculos indicaban justamente lo contrario.

Éste y otros problemas planteados por de Méré motivaron un intercambio de cartas entre Pascal y Fermat en las que por primera vez se formularon los principios fundamentales del Cálculo de probabilidades. Si bien unos pocos problemas sobre juegos de azar habían sido resueltos por matemáticos italianos en los siglos XV y XVI, no existía una teoría general antes de esa famosa correspondencia.

El científico holandés Christian Huygens, maestro de Leibniz, enterado de esa correspondencia publicó rápidamente (en 1657) el primer libro de probabilidades; titulado *De Ratiociniis in Ludo Aleae,* fue un tratado de problemas relacionados con los juegos. El Cálculo de probabilidades llegó a ser pronto popular por sus alusiones a los juegos de azar, y se desarrolló rápidamente a lo largo del siglo XVIII. Quienes más contribuyeron a su desarrollo en ese período fueron Jakob Bernoulli (*) (1654-1705) y Abraham de Moivre (1667-1754).

En 1812, Pierre de Laplace (1749-1827) introdujo gran cantidad de ideas nuevas y técnicas matemáticas en su libro, *Théorie analytique des probabilités.* An-

(*) Citado algunas veces como James Bernoulli.

tes de Laplace, el Cálculo de probabilidades prácticamente consistía en un análisis matemático de los juegos de azar. Laplace demostró que esa teoría podía ser aplicada a multitud de problemas científicos y prácticos. Ejemplos de tales aplicaciones son la *teoría de errores*, la *matemática actuarial* y la *mecánica estadística* que se desarrollaron en el siglo XIX.

Al igual que ha ocurrido con otras muchas ramas de la Matemática, el desarrollo del Cálculo de probabilidades ha sido estimulado por la variedad de sus aplicaciones. Inversamente, cada avance en la teoría ha ampliado el campo de su influencia. La Estadística Matemática es una rama importante del Cálculo de probabilidades aplicado; otras aplicaciones las tenemos en campos tan distintos como la Genética, la Psicología, la Economía y la Ingeniería. Muchos autores han contribuido al desarrollo de la teoría desde el tiempo de Laplace; entre los más importantes están Chebyshev, Markov, von Mises y Kolmogorov.

Una de las dificultades que se presentaron al desarrollar una teoría matemática de la teoría de la probabilidad ha sido alcanzar una definición de probabilidad lo bastante precisa para su utilización matemática, pero lo bastante amplia para que sea aplicable a un número de fenómenos lo mayor posible. La búsqueda de una definición completamente aceptable duró cerca de tres siglos y fue caracterizada por gran número de controversias. El asunto fue definitivamente resuelto en el siglo XX al tratar la teoría de la probabilidad en forma axiomática. En 1933 una monografía del matemático ruso A. Kolmogorov estableció una introducción axiomática que constituyó la base para la moderna teoría. (La traducción inglesa de la monografía de Kolmogorov se titula *Foundations of Probability Theory*, Chelsea, New York, 1950.) Desde entonces las ideas se han ido afinando algo más y hoy día la teoría de la probabilidad es parte de una disciplina más general que es la teoría de la medida.

Este capítulo presenta las nociones fundamentales de la moderna teoría de la probabilidad elemental junto con sus conexiones con la teoría de la medida. También se dan algunas aplicaciones, especialmente a los juegos de azar tales como lanzamiento de monedas, dados y juegos de naipes. Esta exposición pretende poner de manifiesto la estructura lógica del tema como ciencia deductiva y suscitar en el lector el interés en el modo de pensar probabilístico.

13.2 Funciones de conjunto con aditividad finita

El área de una región, la longitud de una curva, o la masa de un sistema de partículas son números que miden la magnitud o contenido de un conjunto. Todas esas medidas tienen ciertas propiedades comunes. Establecidas en forma abstracta, conducen a un concepto general llamado *función de conjunto con aditividad finita*. Más adelante introduciremos la probabilidad como otro ejemplo de una función de este tipo. Para preparar el camino, discutimos primero algunas propiedades comunes a todas esas funciones.

Una función $f: \mathscr{A} \to \mathbf{R}$ cuyo dominio es una colección $\mathscr{A}$ de conjuntos y cuyos valores son números reales, se llama *función de conjunto.* Si A es un conjunto de la colección $\mathscr{A}$, el valor de la función en A se representa por $f(A)$.

DEFINICIÓN DE FUNCIÓN DE CONJUNTO CON ADITIVIDAD FINITA. *Una función de conjunto* $f: \mathscr{A} \to \mathbf{R}$ *se dice que es de aditividad finita si*

$$f(A \cup B) = f(A) + f(B) \tag{13.1}$$

siempre que A *y* B *sean conjuntos disjuntos en* $\mathscr{A}$ *tales que* $A \cup B$ *pertenezca también a* $\mathscr{A}$.

El área, la longitud y la masa son ejemplos de funciones de conjuntos con aditividad finita. En esta sección se discuten consecuencias de la ecuación (13.1).

En las aplicaciones corrientes, los conjuntos de $\mathscr{A}$ son subconjuntos de un conjunto dado S, llamado conjunto universal. Es frecuente tener que efectuar las operaciones de reunión, intersección y complementación sobre los conjuntos de $\mathscr{A}$. Para asegurarse de que $\mathscr{A}$ es cerrado respecto a esas operaciones imponemos a $\mathscr{A}$ que sea un *álgebra booleana,* que se define como sigue.

DEFINICIÓN DE UN ÁLGEBRA BOOLEANA DE CONJUNTOS. *Una clase no vacía* $\mathscr{A}$ *de subconjuntos de un conjunto universal* S *se llama álgebra booleana si para todo par* A *y* B *de conjuntos de* $\mathscr{A}$ *tenemos*

$$A \cup B \in \mathscr{A} \qquad \text{y} \qquad A' \in \mathscr{A}.$$

Aquí $A' = S - A$, *es el complemento de* A *respecto a* S.

Un álgebra booleana $\mathscr{A}$ también es cerrada para las intersecciones y diferencias, ya que tenemos

$$A \cap B = (A' \cup B')' \qquad \text{y} \qquad A - B = A \cap B'.$$

Esto implica que el conjunto vacío $\varnothing$ pertenece a $\mathscr{A}$ ya que $\varnothing = A - A$ para algún A de $\mathscr{A}$. También el conjunto universal S pertenece a $\mathscr{A}$ puesto que $S = \varnothing'$.

A partir de los subconjuntos de un conjunto universal dado S pueden construirse gran número de álgebras booleanas. La menor de esas álgebras es la clase $\mathscr{A}_0 = \{\varnothing, S\}$, que consta tan sólo de dos subconjuntos especiales: $\varnothing$ y S. En el otro extremo está la clase $\mathscr{A}_1$, que consta de *todos* los subconjuntos de S. Toda álgebra booleana $\mathscr{A}$ constituida con subconjuntos de S satisface las relaciones de inclusión $\mathscr{A}_0 \subseteq \mathscr{A} \subseteq \mathscr{A}_1$.

La propiedad de la aditividad finita de funciones de conjunto de la ecuación (13.1) exige que A y B sean conjuntos disjuntos. De esta exigencia se desprende el teorema siguiente.

TEOREMA 13.1. *Si* $f: \mathscr{A} \to \mathbf{R}$ *es una función de conjunto con aditividad finita definida sobre un álgebra de Boole* $\mathscr{A}$ *de conjuntos, entonces para todo par de conjuntos* A *y* B *de* $\mathscr{A}$ *tenemos*

(13.2) $$f(A \cup B) = f(A) + f(B - A),$$

y

(13.3) $$f(A \cup B) = f(A) + f(B) - f(A \cap B).$$

Demostración. Los conjuntos A y $B - A$ son disjuntos y su reunión es $A \cup B$. Luego, aplicando (13.1) a A y a $B - A$ obtenemos (13.2).

Para demostrar (13.3) observemos primero que $A \cap B'$ y B son conjuntos disjuntos cuya reunión es $A \cup B$. Por tanto según (13.1) tenemos

(13.4) $$f(A \cup B) = f(A \cap B') + f(B).$$

Asimismo, $A \cap B'$ y $A \cap B$ son conjuntos disjuntos cuya reunión es A, con lo que (13.1) nos da

(13.5) $$f(A) = f(A \cap B') + f(A \cap B).$$

Restando (13.5) de (13.4) obtenemos (13.3).

13.3 Medidas con aditividad finita

Las funciones de conjunto que representan áreas, longitudes y masas poseen propiedades comunes. Por ejemplo, son todas funciones de conjunto *no negativas.* Esto es,

$$f(A) \geq 0$$

para cada conjunto A de la clase $\mathscr{A}$ que se considera.

DEFINICIÓN DE MEDIDA CON ADITIVIDAD FINITA. *Una función de conjunto no negativa* $f: \mathscr{A} \to \mathbf{R}$ *que es con aditividad finita se dice que es una medida con aditividad finita, o simplemente una medida.*

Aplicando el teorema 13.1 obtenemos inmediatamente las siguientes propiedades de las medidas.

TEOREMA 13.2. *Sea* f: $\mathscr{A} \to \mathbf{R}$ *una medida con aditividad finita definida sobre un álgebra booleana* $\mathscr{A}$. *Para todos los conjuntos* A *y* B *de* $\mathscr{A}$ *tenemos*

a) $f(A \cup B) \leq f(A) + f(B)$.
b) $f(B - A) = f(B) - f(A)$ *si* $A \subseteq B$.
c) $f(A) \leq f(B)$ *si* $A \subseteq B$. (*Propiedad de monotonía*)
d) $f(\varnothing) = 0$.

Demostración. La parte a) se deduce de (13.3), y la parte b) de (13.2). La parte c) resulta de b), y d) se obtiene haciendo $A = B = \varnothing$ en b).

EJEMPLO. *Número de elementos de un conjunto finito.* Sea $S = \{a_1, a_2, \ldots, a_n\}$ un conjunto que consta de n elementos distintos, y sea $\mathscr{A}$ la clase de todos los subconjuntos de S. Para cada A de $\mathscr{A}$, representemos por $\nu(A)$ el número de elementos distintos de A (ν es la letra griega «ni»). Es fácil verificar que esta función es de aditividad finita en $\mathscr{A}$. En efecto, si A tiene k elementos y B m elementos, $\nu(A) = k$ y $\nu(B) = m$. Si A y B son disjuntos es evidente que $A \cup B$ es un subconjunto de S con $k + m$ elementos, así que

$$\nu(A \cup B) = k + m = \nu(A) + \nu(B).$$

Esta función de conjunto es no negativa, con lo que ν es una medida.

13.4 Ejercicios

1. Representemos por $\mathscr{A}$ la clase de todos los subconjuntos de un conjunto universal dado y sean A y B conjuntos cualesquiera de $\mathscr{A}$. Demostrar que:
   a) $A \cap B'$ y B son disjuntos.
   b) $A \cup B = (A \cap B') \cup B$. (Esta fórmula expresa $A \cup B$ como reunión de dos conjuntos disjuntos.)
   c) $A \cap B$ y $A \cap B'$ son disjuntos.
   d) $(A \cap B) \cup (A \cap B') = A$. (Esta fórmula expresa A como reunión de dos conjuntos disjuntos.)
2. El ejercicio 1 b) nos muestra la posibilidad de expresar la reunión de dos conjuntos como reunión de dos conjuntos *disjuntos*. Expresar de manera parecida la reunión de tres conjuntos $A_1 \cup A_2 \cup A_3$ y, más general, de n conjuntos $A_1 \cup A_2 \cup \cdots \cup A_n$. Ilustrar con un diagrama el caso $n = 3$.
3. Al estudiar el conjunto S que consta de 1000 graduados universitarios diez años después de su graduación, se observó que los «triunfantes» formaban un subconjunto A de 400 miembros, los graduados por Caltech otro B de 300, y la intersección $A \cap B$ constaba de 200.
   a) Empleando la notación de la teoría de conjuntos y los conceptos de reunión e intersección de A, B y sus complementarios A' y B' respecto de S, expresar los subconjuntos de aquellas personas de S que poseen las características siguientes:
      I) Ni «triunfantes» ni graduados por Caltech.
      II) «Triunfantes» pero no graduados por Caltech.

iii) «Triunfantes» o graduados por Caltech o ambas cosas.
iv) «Triunfantes» o graduados por Caltech pero no ambas cosas.
v) Pertenecen tan sólo a uno de los subconjuntos A o B.

b) Determinar el número exacto de individuos de cada uno de los cinco subconjuntos anteriores.

4. Sea f una función de conjunto de aditividad finita definida en una clase de conjuntos $\mathscr{A}$. Sean $A_1, \ldots, A_n$ n conjuntos de $\mathscr{A}$ tales que $A_i \cap A_j = \varnothing$ si $i \neq j$. (Una tal colección se denomina *colección disjunta* de conjuntos.) Si la reunión $\bigcup_{k=1}^{m} A_k$ pertenece a $\mathscr{A}$ para todo $m \leq n$, utilizar el método de inducción para demostrar que

$$f\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} f(A_k).$$

En los ejercicios 5, 6 y 7, S representa un conjunto finito que consta de n elementos distintos, sea $S = \{a_1, a_2, \ldots, a_n\}$.

5. Sea $A_1 = \{a_1\}$, el subconjunto que consta del único elemento a_1. Demostrar que la clase $\mathscr{B}_1 = \{\varnothing, A_1, A_1', S\}$ es la más pequeña álgebra de Boole que contiene A_1.
6. Sean $A_1 = \{a_1\}$, $A_2 = \{a_2\}$. En forma parecida a la utilizada en el ejercicio 5, describir la más pequeña álgebra de Boole $\mathscr{B}_2$ que contenga A_1 y A_2.
7. Hacer lo mismo que en el ejercicio 6 para los subconjuntos $A_1 = \{a_1\}$, $A_2 = \{a_2\}$ y $A_3 = \{a_3\}$.
8. Si $\mathscr{B}_k$ representa la mínima álgebra de Boole que contiene los k subconjuntos $A_1 = \{a_1\}$, $A_2 = \{a_2\}, \ldots, A_k = \{a_k\}$, demostrar que $\mathscr{B}_k$ contiene 2^{k+1} subconjuntos de S si $k < n$ y 2^n subconjuntos si $k = n$.
9. Sea f una función de conjunto con aditividad finita definida sobre el álgebra booleana de todos los subconjuntos de un conjunto universal dado S. Supongamos que

$$f(A \cap B) = f(A)f(B)$$

para dos subconjuntos particulares A y B de S. Si $f(S) = 2$, demostrar que

$$f(A \cup B) = f(A') + f(B') - f(A')f(B').$$

10. Si A y B son dos conjuntos, su *diferencia simétrica* $A \triangle B$ es el conjunto definido por $A \triangle B = (A - B) \cup (B - A)$. Demostrar cada una de las siguientes propiedades de la diferencia simétrica.
    a) $A \triangle B = B \triangle A$.
    b) $A \triangle A = \varnothing$.
    c) $A \triangle B \subseteq (A \triangle C) \cup (C \triangle B)$.
    d) $A \triangle B$ es disjunto con la intersección de A y B.
    e) $(A \triangle B) \triangle C = A \triangle (B \triangle C)$.
    f) Si f es una función de conjunto con aditividad finita definida sobre el álgebra de Boole $\mathscr{A}$ de todos los subconjuntos de un conjunto dado S, entonces para todo par A y B de $\mathscr{A}$ tenemos $f(A \triangle B) = f(A) + f(B) - 2f(A \cap B)$.

13.5 Definición de probabilidad para espacios muestrales finitos

En el lenguaje de las funciones de conjunto, la probabilidad es un tipo especial de medida (representada aquí por P) definida sobre una particular álgebra booleana $\mathscr{B}$ de conjuntos. Los elementos de $\mathscr{B}$ son subconjuntos de un conjunto universal S. En la teoría de la probabilidad el conjunto universal S se llama *espacio muestral.* Primero comentaremos la definición de probabilidad para espacios muestrales finitos y luego lo haremos para los infinitos.

DEFINICIÓN DE PROBABILIDAD PARA ESPACIOS MUESTRALES FINITOS. *Sea $\mathscr{B}$ un álgebra de Boole cuyos elementos son subconjuntos de un conjunto finito dado S. Una función de conjunto P definida en $\mathscr{B}$ se llama medida de probabilidad si satisface las tres condiciones siguientes:*

a) *P es de aditividad finita.*
b) *P es no negativa.*
c) $P(S) = 1$.

Dicho de otro modo, para los espacios muestrales finitos la probabilidad es simplemente una medida que asigna el valor 1 al espacio completo.

Importa darse cuenta de que para una descripción completa de la medida de probabilidad deben precisarse tres ideas: el espacio muestral S, el álgebra de Boole $\mathscr{B}$ constituida con ciertos subconjuntos de S, y la función de conjunto P. La terna $(S, \mathscr{B}, P)$ se denomina frecuentemente *espacio de probabilidad.* En la mayoría de las aplicaciones elementales el álgebra de Boole $\mathscr{B}$ es la colección de *todos* los subconjuntos de S.

EJEMPLO. El juego de «cara o cruz» es una aplicación de la teoría de probabilidad. Como espacio muestral S tomamos el conjunto de todos los resultados posibles en el juego. Cada resultado es o «cara» o «cruz» que representamos con los símbolos h y t. Dicho espacio muestral es pues $\{h, t\}$ es decir, el conjunto que consta de h y t. Como álgebra booleana consideramos la colección de todos los subconjuntos de S; éstos son cuatro, $\varnothing$, S, H y T, donde $H = \{h\}$ y $T = \{t\}$. Asignemos ahora probabilidades a cada uno de esos subconjuntos. Para $\varnothing$ y S no tenemos opción a elegir valores de probabilidad. Por la propiedad b), $P(S) = 1$, y como *P es no negativa,* $P(\varnothing) = 0$. En cambio, tenemos libertad en la asignación de la probabilidad a los otros dos subconjuntos, H y T. Ya que H y T son conjuntos disjuntos cuya reunión es S, la propiedad aditiva exige que

$$P(H) + P(T) = P(S) = 1 .$$

Como valores de $P(H)$ y $P(T)$ podemos tomar valores cualesquiera no negativos con tal de que su suma sea 1. Si tenemos en cuenta que la moneda es imparcial

de modo que no existe razón a *priori* para preferir cara o cruz, parece natural asignar los valores

$$P(H) = P(T) = \tfrac{1}{2}.$$

Si, en cambio, la moneda no es geométricamente perfecta, podremos asignar valores diferentes a esas dos probabilidades. Por ejemplo, los valores $P(H) = \frac{1}{3}$ y $P(T) = \frac{2}{3}$ son tan aceptables como $P(H) = P(T) = \frac{1}{2}$. En efecto, para todo número real p en el intervalo $0 \leqslant p \leqslant 1$ podemos definir $P(H) = p$ y $P(T) = 1 - p$, y la función resultante P satisfará todas las condiciones que se exigen a una medida de probabilidad.

Para una moneda determinada, no existe un método matemático para precisar cuál es la probabilidad p «real». Si escogemos $p = \frac{1}{2}$ podemos deducir consecuencias lógicas en la hipótesis de que la moneda es correcta y no presenta sesgo. La teoría desarrollada para el estudio de las probabilidades en monedas correctas, puede utilizarse como test comprobatorio de su carencia de sesgo, efectuando un gran número de experiencias con ella y comparando los resultados experimentales con las predicciones teóricas. El poner de acuerdo la teoría y la evidencia empírica pertenece a la rama de aplicación de la teoría de la probabilidad que se llama *inferencia estadística,* y que no expondremos en este libro.

El ejemplo anterior es una típica aplicación del Cálculo de probabilidades. Las cuestiones probabilísticas se presentan a menudo en situaciones llamadas «experimentos». No intentaremos definir un experimento; en cambio, mencionaremos tan sólo algunos ejemplos corrientes: lanzar una o varias monedas, echar un par de dados, repartir una mano de bridge, sacar una bola de una urna, recuento de las muchachas estudiantes en el Instituto Tecnológico de California, selección de un número en una guía telefónica, registro de la radiación en un contador Geiger.

Para discutir las cuestiones de probabilidad que surgen en tales experimentos, nuestro primer trabajo es la construcción de un espacio muestra S que pueda utilizarse para representar todos los resultados posibles del experimento, como hicimos en el juego de cara y cruz. Cada elemento de S representará un resultado del experimento y cada resultado corresponderá a uno y sólo un elemento de S. A continuación, elegimos un álgebra de Boole $\mathscr{B}$ de subconjuntos de S (casi siempre *todos* los subconjuntos de S) y entonces se define una medida de probabilidad P sobre $\mathscr{B}$. La elección de $S, \mathscr{B}$, y P dependerá de la información que se posea acerca de los detalles del experimento y del problema que nos vamos a plantear. El objeto del Cálculo de probabilidades no es discutir si el espacio de probabilidad $(S, \mathscr{B}, P)$ ha sido elegido correctamente. Esto pertenece a la ciencia o juego del que el experimento ha surgido, y tan sólo la experiencia puede darnos idea de si la elección fue bien hecha o no lo fue. *El Cálculo de probabilidades es el estudio de las conse-*

cuencias lógicas que pueden deducirse una vez está dado el espacio de probabilidad. La elección de un buen espacio de probabilidad no es teoría de probabilidad — ni siquiera es matemática; es, en cambio, parte del arte de aplicar la teoría probabilística al mundo real.

Si $S = \{a_1, a_2, \ldots, a_n\}$, y si $\mathscr{B}$ consta de todos los subconjuntos de S, la función de probabilidad P está completamente determinada si conocemos sus valores para los conjuntos de un solo elemento,

$$P(\{a_1\}), P(\{a_2\}), \ldots, P(\{a_n\}).$$

En efecto, todo subconjunto A de S es una reunión disjunta de los conjuntos anteriores, y $P(A)$ está determinada por la propiedad aditiva. Por ejemplo, cuando

$$A = \{a_1\} \cup \{a_2\} \cup \cdots \cup \{a_k\},$$

la propiedad aditiva exige que

$$P(A) = \sum_{i=1}^{k} P(\{a_i\}).$$

Para simplificar la notación y la terminología, escribimos $P(a_i)$ en lugar de $P(\{a_i\})$. Este número también se llama *probabilidad del punto* a_i. Por lo tanto, la asignación de la probabilidad puntual $P(x)$ a cada elemento x de un conjunto finito S equivale a una descripción completa de la función de probabilidad P.

13.6 Terminología propia del cálculo de probabilidades

Cuando se habla de probabilidades, a menudo se oyen frases tales como «dos sucesos son igualmente probables», «un suceso es imposible», o «un suceso es cierto». Las expresiones de este tipo tienen sentido intuitivo y es agradable y útil saber emplear un lenguaje tan lleno de colorido en las discusiones matemáticas. Antes de hacerlo así, no obstante, es necesario exponer el significado de este lenguaje usando los conceptos fundamentales de nuestra teoría.

Debido a que el método probabilístico se usa en cuestiones prácticas, es conveniente imaginarse que cada espacio de probabilidad $(S, \mathscr{B}, P)$ está asociado a un experimento real o ideal. El conjunto universal S puede entonces concebirse como la colección de todos los resultados imaginables del experimento, como en el ejemplo de la moneda comentado en la sección precedente. Cada elemento de S se llama *resultado* o *muestra* y los subconjuntos de S que se presentan en el álgebra de Boole $\mathscr{B}$ se denominan *sucesos.* Los motivos de esta terminología se pondrán en evidencia al tratar algunos ejemplos.

Supongamos un espacio de probabilidad $(S, \mathscr{B}, P)$ asociado a un experimento. Sea A un suceso, y supongamos que el experimento se lleva a cabo y que su resultado es x. (En otras palabras, sea x un punto de S.) Este resultado puede o no

pertenecer al conjunto A. Si pertenece, se dice *que el suceso A ha ocurrido*. En el caso contrario, *el suceso A no ha ocurrido*, en cuyo caso $x \in A'$, lo que equivale a decir que ha ocurrido el suceso complementario A'. Un suceso A *es imposible* si $A = \varnothing$, porque en este caso ningún resultado del experimento puede pertenecer a A. El suceso A es cierto si $A = S$, porque entonces todo resultado de la prueba pertenece a A.

La función de probabilidad P asigna a cada suceso A una probabilidad $P(A)$. [Del valor de $P(A)$ y del modo de asignarlo no nos ocuparemos por el momento.] El número $P(A)$ se llama *la probabilidad de que un resultado de la prueba sea un elemento de A*. También decimos que $P(A)$ es *la probabilidad de que el suceso A ocurra al efectuar* el experimento.

Al suceso imposible $\varnothing$ debe asignarse probabilidad cero porque P es una medida de aditividad finita. No obstante, existen sucesos con probabilidad cero y que no son imposibles. En otras palabras, algunos de los subconjuntos no vacíos de S pueden tener asignada probabilidad cero. Al suceso cierto S se asigna probabilidad 1 según la correcta definición de probabilidad, pero pueden existir otros subconjuntos que también tienen asignada probabilidad 1. En el ejemplo 1 de la sección 13.8 se citan conjuntos no vacíos con probabilidad cero y subconjuntos propios de S que tienen probabilidad 1.

Dos sucesos A y B son *igualmente probables* si $P(A) = P(B)$. El suceso A es *más probable* que el B si $P(A) > P(B)$, y *por lo menos tan probable* como el B si $P(A) \geqslant P(B)$. La tabla 13.1 nos muestra una lista de locuciones del lenguaje habitual en las discusiones de la teoría de probabilidades. Las letras A y B representan sucesos, y x el resultado de un experimento asociado al espacio muestral S. Cada fila de la columna de la izquierda es una afirmación relativa a los sucesos A y B, y en la misma fila en la columna de la derecha se expresa la misma afirmación en el lenguaje de la teoría de conjuntos.

TABLA 13.1 Proposiciones usadas en la Teoría de Probabilidades y su significado en la de conjuntos

Proposiciones	*Significado en la teoría de conjuntos*
Por lo menos uno de los sucesos A o B ocurre	$x \in A \cup B$
Ambos sucesos A y B ocurren	$x \in A \cap B$
Ni A ni B ocurren	$x \in A' \cap B'$
A ocurre y B no	$x \in A \cap B'$
Exactamente ocurre uno de los sucesos A o B	$x \in (A \cap B') \cup (A' \cap B)$
No más de uno de los sucesos A o B ocurre	$x \in (A \cap B)'$
Si A ocurre, también B (A implica B)	$A \subseteq B$
A y B se excluyen mutuamente	$A \cap B = \varnothing$
Suceso A o suceso B	$A \cup B$
Suceso A y suceso B	$A \cap B$

13.7 Ejercicios

Sean S un espacio muestral dado y A, B, C sucesos cualesquiera (esto es, subconjuntos de S en la correspondiente álgebra de Boole $\mathscr{B}$). Cada una de las afirmaciones de los ejercicios del 1 al 12 se expresa con una proposición. Expresar dichas afirmaciones como reuniones e intersecciones de A, B y C y de sus complementos.

1. Si ocurre A, no ocurre B.
2. Ninguno de los sucesos A, B, C ocurre.
3. Tan sólo ocurre A.
4. Por lo menos uno de los A, B, C ocurre.
5. Uno exactamente de los A, B, C ocurre.
6. No ocurre más que uno.
7. Por lo menos dos de los A, B, C ocurren.
8. Ocurren exactamente dos.
9. Ocurren no más de dos.
10. Ocurren A y C, pero no B.
11. Ocurren los tres sucesos.
12. Ocurren no más de tres.

13. A designa el suceso de conseguir un total impar al lanzar dos dados, y B el suceso de sacar por lo menos un 6. Expresar por una frase cada uno de los sucesos siguientes:

a) $A \cup B$,
b) $A \cap B$,
c) $A \cap B'$,
d) $A' \cap B$,
e) $A' \cap B'$,
f) $A' \cup B$.

14. Sean A y B dos sucesos. Demostrar que

$$P(A \cap B) \leq P(A) \leq P(A \cup B) \leq P(A) + P(B).$$

15. A y B representan dos sucesos y sean $a = P(A)$, $b = P(B)$, $c = P(A \cap B)$. Calcular en función de a, b, y c, las probabilidades de los sucesos siguientes:

a) A',
b) B,
c) $A \cup B$,
d) $A' \cup B'$,
e) $A' \cup B$,
f) $A \cap B'$.

16. Dados tres sucesos A, B, C. Demostrar que

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

13.8 Ejemplos resueltos

A continuación veremos cómo pueden usarse los conceptos de las secciones precedentes, para resolver problemas típicos de probabilidades.

EJEMPLO 1. ¿Cuál es la probabilidad de que por lo menos una cara salga en dos tiradas de una moneda?

Primera solución. El experimento consiste en lanzar una moneda dos veces; el conjunto S de todos los resultados posibles puede expresarse como sigue:

$$S = \{hh, ht, th, tt\}.$$

Si aceptamos que esos resultados son igualmente probables, asignamos la probabilidad puntual $P(x) = \frac{1}{4}$ a cada x de S. El suceso «que se presente por lo menos una cara» puede representarse por el subconjunto

$$A = \{hh, ht, th\}.$$

La probabilidad de este suceso es la suma de las probabilidades puntuales de sus elementos. Luego, $P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$.

Segunda solución. Supongamos que se usa el mismo espacio muestral pero que asignamos las probabilidades puntuales como sigue: (*)

$$P(hh) = 1, \qquad P(ht) = P(th) = P(tt) = 0.$$

Entonces la probabilidad del suceso «por lo menos sale una cara» es

$$P(hh) + P(ht) + P(th) = 1 + 0 + 0 = 1.$$

El hecho de que lleguemos a un resultado final distinto del conseguido en la primera solución no debe alarmar al lector. Partimos de un conjunto de premisas distinto. Consideraciones de tipo psicológico pueden llevarnos a pensar que la asignación de probabilidades en la primera solución es la más natural. No obstante, muchos podrían estar de acuerdo en que esto es así si la moneda es «imparcial». Sin embargo, si la moneda está «cargada» de manera que siempre salga cara, la asignación hecha en la segunda solución es más natural.

El ejemplo anterior nos demuestra que no cabe esperar respuesta única a la pregunta formulada en el mismo. Para contestar en forma adecuada tenemos que especificar la elección del espacio muestral y la asignación de las probabilidades puntuales. La probabilidad de un suceso puede deducirse de manera lógica únicamente cuando se conocen el espacio muestral y la asignación de las probabilidades puntuales. Eligiendo de distintos modos el espacio muestral y las probabilidades puntuales puede llegarse a contestaciones «correctas» diferentes de la misma cuestión.

A veces, la asignación de probabilidades a los resultados particulares de un

(*) Nótese que para esta asignación de probabilidades hay subconjuntos de S no vacíos con probabilidad cero y subconjuntos propios con probabilidad 1.

experimento viene dictada por el lenguaje utilizado al describirlo. Por ejemplo, cuando un objeto es elegido «al azar» en un conjunto finito de n elementos, se entiende que cada resultado es igualmente probable y podría asignársele probabilidad $1/n$. Análogamente, cuando lanzamos una moneda o un dado, si no tenemos a priori motivo alguno para pensar que la moneda o el dado no son perfectos, suponemos que todos los resultados son igualmente probables. Este convenio se adoptará en todos los ejercicios de este capítulo.

EJEMPLO 2. Si un naipe es sacado al azar de cada una de dos barajas, ¿cuál es la probabilidad de que por lo menos uno sea el as de corazones?

Solución. El experimento consiste en tomar dos naipes, a y b, uno de cada baraja. Representaremos un resultado por un par ordenado (a, b). El número de resultados posibles, esto es, el número total de pares distintos (a, b) del espacio muestral S es 52^2. Asignamos a cada uno de esos pares la probabilidad $1/52^2$. El suceso en el que estamos interesados es el conjunto A de pares (a, b), en los que o a o b es el as de corazones. En A hay $52 + 51$ elementos. Por tanto, en estas hipótesis deducimos que

$$P(A) = \frac{52 + 51}{52^2} = \frac{1}{26} - \frac{1}{52^2}.$$

EJEMPLO 3. Si dos naipes se sacan al azar de una baraja; ¿cuál es la probabilidad de que una de ellas sea un as de corazones?

Solución. Como en el ejemplo 2 empleamos pares ordenados (a, b) como elementos del espacio muestral. En este caso el citado espacio tiene $52 \cdot 51$ elementos y el suceso A que se considera tiene $51 + 51$ elementos. Si asignamos la probabilidad puntual $1/(52 \cdot 51)$ a cada resultado obtenemos

$$P(A) = \frac{2 \cdot 51}{52 \cdot 51} = \frac{1}{26}.$$

EJEMPLO 4. ¿Cuál es la probabilidad de sacar 6 o menos de 6 con tres dados.

Solución. Designamos cada resultado del experimento como una terna (a, b, c) donde a, b y c pueden tomar valores de 1 a 6. Por lo tanto, el espacio muestral consta de 6^3 elementos y asignamos la probabilidad $1/6^3$ a cada resultado. El suceso A en cuestión es el conjunto de todas las ternas que satisfacen la desigualdad $3 \leq a + b + c \leq 6$. Si A_n representa los conjuntos (a, b, c) para los cuales $a + b + c = n$, tenemos

$$A = A_3 \cup A_4 \cup A_5 \cup A_6.$$

La enumeración directa muestra que los conjuntos A_n, con $n = 3, 4, 5$ y 6 contienen 1, 3, 6 y 10 elementos respectivamente. Por ejemplo, el conjunto A_6 viene dado por

$$A_6 = \{(1, 2, 3), (1, 3, 2), (1, 1, 4), (1, 4, 1), (2, 1, 3),$$
$$(2, 3, 1), (2, 2, 2), (3, 1, 2), (3, 2, 1), (4, 1, 1)\}.$$

Por tanto, A tiene 20 elementos y

$$P(A) = \frac{20}{6^3} = \frac{5}{54}.$$

EJEMPLO 5. Se lanza un dado una vez. ¿Cuál es la probabilidad de que el número de puntos conseguido sea par o múltiplo de 3?

Solución. Elegimos el espacio muestral $S = \{1, 2, 3, 4, 5, 6\}$, que consta de seis elementos, a cada uno de los cuales asignamos la probabilidad $1/6$. El suceso «par» es el conjunto $A = \{2, 4, 6\}$, el suceso «múltiplo de 3» es $B = \{3, 6\}$. Nos interesa su reunión, que es el conjunto $A \cup B = \{2, 3, 4, 6\}$. Puesto que este conjunto posee cuatro elementos tenemos $P(A \cup B) = 4/6$.

Este ejemplo puede resolverse de otro modo, usando la fórmula

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \tfrac{3}{6} + \tfrac{2}{6} - \tfrac{1}{6}.$$

13.9 Ejercicios

1. Sea S un espacio muestral finito de n elementos. Supongamos que asignamos la misma probabilidad a cada uno de los puntos de S. Sea A un subconjunto de S que conste de k elementos. Demostrar que $P(A) = k/n$.

En cada uno de los ejercicios 2 al 8, describir la elección del espacio muestral y la asignación de probabilidades. En los ejercicios relacionados con juegos de naipes, se supone que todos los naipes tienen la misma probabilidad de ser repartidos.

2. Se mezclan cinco monedas falsas con nueve auténticas.
   a) Se selecciona al azar una moneda. Calcular la probabilidad de que sea falsa.
   Si se seleccionan dos monedas, calcular la probabilidad de que:
   b) una sea buena y una falsa.
   c) las dos sean falsas.
   d) las dos sean buenas.
3. Calcular la probabilidad de cada uno de los sucesos que se describieron en el ejercicio 13 de la sección 13.7. Asignar la misma probabilidad a cada uno de los 36 elementos del espacio muestral.

4. ¿Cuál es la probabilidad de sacar por lo menos uno de los resultados 7, 11, o 12 con dos dados?
5. Una mano de poker contiene cuatro corazones y una espada. La espada se rechaza y se toma un naipe de los restantes de la baraja. Calcular la probabilidad de sacar un quinto corazón.
6. En el poker, una escalera son cinco cartas consecutivas no necesariamente del mismo palo. Si una mano contiene cuatro cartas consecutivas (pero no A234 o JQKA) y una quinta carta no consecutiva a las otras, calcular la probabilidad de lograr la escalera, al devolver la quinta carta y pedir una nueva.
7. Una mano de poker contiene cuatro de las cinco cartas en sucesión pero con un «salto» intercalado (por ejemplo, 5689), la quinta carta se devuelve y se toma otra nueva del resto de la baraja. Calcular la probabilidad de conseguir llenar el «hueco intermedio».
8. Una urna contiene A bolas blancas y B bolas negras. Una segunda urna contiene C bolas blancas y D bolas negras. Se saca al azar un bola de la primera urna y se echa en la segunda urna. Seguidamente se extrae al azar una bola de la segunda urna. Calcular la probabilidad de cada uno de los sucesos siguientes:
   a) La primera bola es blanca.
   b) La primera bola es negra.
   c) La segunda bola es blanca, habiendo sido blanca la bola transferida.
   d) La segunda bola es blanca, habiendo sido negra la bola transferida.
9. Se extraen dos bolas de una urna devolviendo la bola después de la primera extracción. La urna contiene cuatro bolas rojas y dos blancas. Calcular la probabilidad de cada uno de los sucesos siguientes.
   a) Ambas bolas son blancas.
   b) Ambas bolas son rojas.
   c) Ambas bolas son del mismo color.
   d) Por lo menos una bola es roja.
10. Sea P_n la probabilidad de que ocurran exactamente n sucesos de los A y B, tomando n los valores 0, 1, 2. Expresar cada uno de los números P_0, P_1, P_2 en función de $P(A)$, $P(B)$ y $P(A \cap B)$.

Puntos a favor. Algunos juegos de azar se expresan en función de la «suerte a favor» o «puntos a favor» mejor que en función de las probabilidades. Por ejemplo, si lanzamos un dado, la probabilidad de sacar un tres es $1/6$. Los resultados posibles son seis, uno de ellos es favorable y cinco desfavorables. Esto a menudo se expresa diciendo que la suerte a favor del suceso está como 1 a 5, o que la suerte en contra es de 5 a 1. Entonces suele relacionarse ésta con la probabilidad mediante la igualdad

$$\frac{1}{6} = \frac{1}{1+5}.$$

En general, si A es un suceso con probabilidad $P(A)$ y si a y b son dos números reales que

$$P(A) = \frac{a}{a+b}, \tag{13.6}$$

decimos que la *suerte a favor de A* es de a a b, o que *la suerte en contra de A* es de b a a. Puesto que $1 - a/(a+b) = b/(a+b)$, la suerte en contra de A es la misma que la suerte a favor del suceso complementario A'. El ejercicio siguiente está dedicado a otras propiedades de este concepto de «los puntos o casos favorables» y sus relaciones con las probabilidades.

11. Si $P(A) = 1$, probar que (13.6) sólo puede satisfacerse si $b = 0$ y $a \neq 0$. Si $P(A) \neq 1$, demostrar que existen infinidad de elecciones posibles de a y b que satisfacen (13.6) pero que todas tienen la misma razón a/b.
12. Calcular los puntos a favor de cada uno de los sucesos descritos en el ejercicio 2.
13. Dados los sucesos A y B. Si la suerte en contra de A es de 2 a 1 y a favor de $A \cup B$ de 3 a 1, demostrar que

$$\tfrac{5}{12} \leq P(B) \leq \tfrac{3}{4}.$$

Dar un ejemplo en el que $P(B) = \frac{5}{12}$ y otro en el que $P(B) = \frac{3}{4}$.

13.10 Algunos principios básicos de análisis combinatorio

Muchos problemas del cálculo de probabilidades y de otras ramas de la Matemática pueden reducirse a problemas de recuento del número de elementos de un conjunto finito. Los métodos sistemáticos para estudiar tales problemas forman parte de una disciplina matemática conocida con el nombre de *análisis combinatorio*. En esta sección nos referimos brevemente a algunas ideas fundamentales de análisis combinatorio que son útiles al analizar algunos de los más complicados problemas de la teoría de la probabilidad.

Si tenemos a la vista todos los elementos de un conjunto finito, no resulta difícil contar el número de los mismos. Sin embargo, muy a menudo, un conjunto se describe de tal manera que es imposible «ver» todos sus elementos. Por ejemplo, deseamos conocer el número de manos de bridge distintas que se pueden repartir. Cada jugador recibe 13 naipes de los 52 de la baraja. El número posible de manos distintas es el mismo que el de subconjuntos distintos de 13 elementos distintos que se pueden formar con un conjunto de 52 elementos también distintos. Ya que este número supera los 635 mil millones, la enumeración directa de todas las posibilidades no es el mejor método de atacar el problema; no obstante, con el Análisis combinatorio puede resolverse sin dificultad.

Este problema es un caso particular del problema más general de calcular el número de subconjuntos distintos de k elementos que pueden formarse con un conjunto de n elementos (*), siendo $n \geqslant k$. Designemos este número por $f(n, k)$. Ya es conocido que

$$f(n, k) = \binom{n}{k}, \tag{13.7}$$

donde, como de costumbre, $\binom{n}{k}$ representa el coeficiente binómico,

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

(*) Cuando decimos que un conjunto tiene n elementos, significamos que consta de n elementos *distintos*.

En el problema de las manos de bridge tenemos $f(52, 13) = \binom{52}{13} = 635\,013\,559\,600$ manos distintas que un jugador puede recibir.

Existen muchos métodos conocidos para llegar a (13.7). El más directo consiste en formar cada subconjunto de k elementos eligiendo éstos uno a uno. Existen n posibilidades en la primera elección, $n - 1$ en la segunda, y $n - (k - 1)$ en la k-ésima. Si efectuamos todas las elecciones posibles de esta manera, obtenemos un total de

$$n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

subconjuntos de k elementos. Naturalmente, no todos esos conjuntos son distintos. Por ejemplo, si $k = 3$ los seis subconjuntos.

$$\{a, b, c\}, \{b, c, a\}, \{c, a, b\}, \{a, c, b\}, \{c, b, a\}, \{b, a, c\}$$

son todos iguales. En general, con este método, cada subconjunto de k elementos se cuenta exactamente $k!$ veces.(*). Por consiguiente, debemos dividir el número $n!/(n - k)$ por $k!$ para obtener $f(n, k)$. Esto nos da $f(n, k) = \binom{n}{k}$, como se afirmó.

Este razonamiento es más o menos característico del Análisis combinatorio que se precisa en las próximas secciones. Por consiguiente, parece natural comentar brevemente los principios fundamentales en los que se apoya este Análisis.

A menudo nos interesa contar el número de elementos del producto cartesiano de n conjuntos finitos $A_1, \ldots, A_n$. El producto cartesiano se representa con el símbolo $A_1 \times \ldots \times A_n$ y se define así:

$$A_1 \times \cdots \times A_n = \{(a_1, \ldots, a_n) \mid a_1 \in A_1, \ldots, a_n \in A_n\}.$$

Esto es, el producto cartesiano consta del conjunto de todas las n-plas ordenadas $(a_1, \ldots, a_n)$ en las que el k-ésimo elemento de cada n-pla procede del k-ésimo conjunto A_k.

En la figura 13.1 se muestra un ejemplo con $n = 2$. En él $A_1 = \{1, 2, 4, 5\}$ y $A_2 = \{1, 3\}$. Existen 4 elementos en A_1 y 2 en A_2, dando un total de 8 elementos en el producto cartesiano $A_1 \times A_2$. Con mayor generalidad, si A_1 consta de k_1 elementos y A_2 de k_2 elementos, entonces $A_1 \times A_2$ constará de k_1k_2 elementos. Por inducción respecto a n, se deduce que si A_r consta de k_r elementos, el producto cartesiano $A_1 \times \ldots \times A_n$ constará de $k_1 \ldots k_n$ elementos.

Para expresar este resultado en la terminología de las funciones de conjunto, sea $\mathscr{F}$ la clase de todos los conjuntos finitos y ν la función de conjunto definida

(*) En el ejemplo 3 en la página 590, se ve con claridad la explicación de esto, y damos una deducción de (13.7) más rigurosa.

sobre $\mathscr{F}$ así: Si $A \in \mathscr{F}$, $\nu(A)$ representa el número de elementos distintos de A. (Para el conjunto vacío definimos $\nu(\varnothing) = 0$.) Es entonces fácil comprobar que ν es una función de conjunto de aditividad finita, podemos pues escribir,

$$\nu\left(\bigcup_{i=1}^{n} S_i\right) = \sum_{i=1}^{n} \nu(S_i) \tag{13.8}$$

si $\{S_1, S_2, \ldots, S_n\}$ es una colección de conjuntos finitos disjuntos (esto es, si $S_i \cap S_j = \varnothing$ siempre que $i \neq j$). El número de elementos de un producto cartesiano puede expresarse en función de ν como sigue:

$$\nu(A_1 \times A_2 \times \cdots \times A_n) = \nu(A_1)\nu(A_2) \cdots \nu(A_n).$$

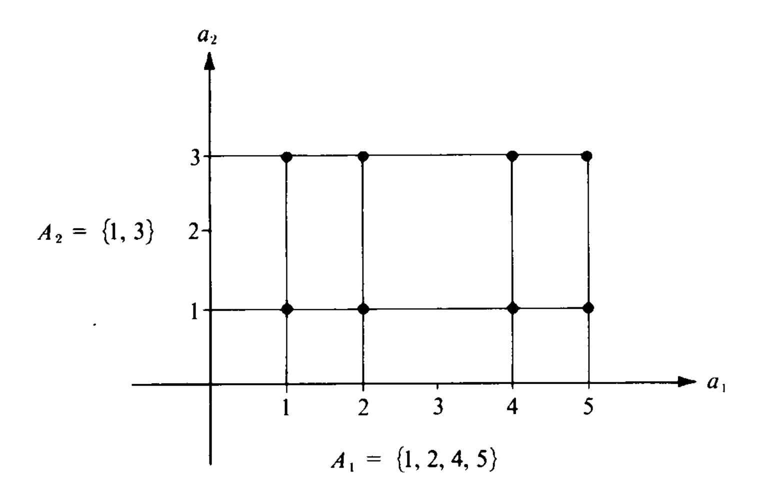


FIGURA 13.1 *Producto cartesiano de dos conjuntos. Los puntos marcados representan* $A_1 \times A_2$.

Una fórmula parecida nos dice cómo contar el número de elementos de cualquier conjunto T de n-plas si conocemos el número de elecciones posibles para cada una de las sucesivas componentes. Por ejemplo, supongamos que existen k_1 elecciones posibles para el primer componente x_1. Sea k_2 el número de elecciones posibles para x_2, una vez x_1 es conocido. Análogamente, k_r sea el número de elecciones posibles para x_r, una vez $x_1, x_2, \ldots, x_{r-1}$ han sido elegidos. El número de n-plas que pueden formarse con estas elecciones es:

$$\nu(T) = k_1 k_2 \cdots k_n .$$

Esta fórmula se denomina frecuentemente *regla del cálculo secuencial.* Puede demostrarse por inducción. En muchas aplicaciones el conjunto de elecciones para x_r puede ser difícil de precisar, pues no puede determinarse hasta después de haber elegido los componentes anteriores. (Este es el caso del cálculo de las manos de bridge.) Afortunadamente, al aplicar la regla del cálculo secuencial no necesitamos conocer ese conjunto de elecciones de x_r, sino únicamente el número de elecciones posibles para x_r.

La propiedad aditiva expresada por la fórmula (13.8) y la regla del cálculo secuencial nos proporcionan la solución de muchos problemas de cálculo. En los ejemplos siguientes vamos a ver el modo de aplicarlos.

EJEMPLO 1. *Muestreo con reposición.* Sea un conjunto S que conste de n elementos. Si $k \geqslant 1$, ¿cuántas k-plas pueden formarse si cada componente puede ser un elemento cualquiera de S?

Observación. Puede ser útil imaginar S como una urna que contenga n bolas numeradas $1, 2, \ldots, n$. Sacamos una bola y tomamos su marca como primer componente de nuestra k-pla. Devolvemos la bola a la urna, volvemos a sacar otra bola y tomamos su marca como segunda componente y así sucesivamente, hasta que hemos completado k extracciones. De este modo, la misma marca puede aparecer varias veces en la k-pla formada.

Solución. Cada k-pla es un elemento del producto cartesiano

$$T = S_1 \times \cdots \times S_k,$$

siendo cada $S_i = S$. Recíprocamente, cada elemento de T es una de las k-plas que consideramos. Luego el número de las k-plas formadas de este modo es

$$\nu(T) = \nu(S_1) \cdots \nu(S_k) = n^k.$$

EJEMPLO 2. *Muestreo sin reposición.* Dado un conjunto S de n elementos. Si $k \leqslant n$, ¿cuántos k-plas pueden formarse si los componentes se eligen en S *sin* reposición, es decir, si ningún elemento de S puede usarse más de una vez en una misma k-pla?

Solución. Para el primer componente x_1 existen n elecciones posibles (los n elementos de S). Una vez elegido x_1, quedan $n - 1$ elecciones posibles para x_2. Elegido x_2, que dan $n - 2$ posibilidades para x_3, y así sucesivamente, para x_k tendremos $n - k + 1$ posibilidades de elección. Por lo tanto, según la regla del cálculo secuencial, el número total de k-plas así formadas es:

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

En particular, cuando $k = n$ resultan $n!$ n-plas distintas.

EJEMPLO 3. *Número de subconjuntos de k elementos de un conjunto dado de n elementos.* Si $k \leqslant n$, ¿cuántos subconjuntos distintos de k elementos pueden formarse a partir de un conjunto dado S de n elementos?

Solución. Designemos por r el número de subconjuntos en cuestión y éstos por

$$A_1, A_2, \ldots, A_r.$$

Estos conjuntos son distintos, pero no necesariamente disjuntos. Calcularemos r en función de n y k por un método indirecto. A tal fin, designemos con B_i la colección de k-plas que pueden formarse eligiendo los componentes entre los elementos de A_i sin reposición. Los conjuntos $B_1, B_2, \ldots, B_r$ *son* disjuntos. Además, si aplicamos el resultado del ejemplo 2 con $n = k$ tenemos:

$$\nu(B_i) = k! \quad \text{para cada } i = 1, 2, \ldots, r.$$

Pongamos ahora

$$T = B_1 \cup B_2 \cup \cdots \cup B_r.$$

Este conjunto T consta de todas las k-plas que se pueden formar eligiendo los componentes en S sin reposición. Del ejemplo 2 tenemos:

$$\nu(T) = n!/(n-k)!$$

y por la aditividad tenemos también,

$$\nu(T) = \sum_{i=1}^{r} \nu(B_i) = k!\, r.$$

Igualando las dos expresiones de $\nu(T)$ obtenemos:

$$r = \frac{n!}{k!\,(n-k)!} = \binom{n}{k}.$$

Esto demuestra la fórmula (13.7) establecida antes en esta sección.

Si utilizamos el resultado del ejemplo 3 para calcular el número total de subconjuntos de un conjunto S que consta de n elementos, obtenemos:

$$\sum_{k=0}^{n} \binom{n}{k}.$$

Ya que esta suma también se obtiene al desarrollar $(1 + 1)^n$ por la fórmula binomial, el número de subconjuntos de S es 2^n.

13.11 Ejercicios

1. Sea $A = \{1, 2, 3\}$. Desarrollar según la notación en lista el conjunto de pares ordenados (a, b) obtenidos eligiendo el primer componente en A y el segundo componente entre los *restantes* elementos de A. Tal conjunto de pares ordenados, ¿puede expresarse en forma de producto cartesiano?
2. Una pareja de naipes puede ser obtenida de una baraja de 52 naipes, de $52 \cdot 51 = 2652$ maneras. Determinar el número de parejas *distintas* y exponer el razonamiento.
3. Un comité del Senado norteamericano que consta de seis demócratas y cuatro republicanos debe elegir un presidente y un vicepresidente. ¿De cuántas maneras se pueden elegir este par de funcionarios si el presidente debe ser demócrata?
4. Un juego consiste en tirar una moneda dos veces y lanzar luego un dado. Expresar cada resultado de este experimento como una terna ordenada (a, b, c), donde cada a y b pueden ser o C (cara) o ✝ (cruz) y c es el número de puntos que marca el dado. por ejemplo $(C, C, 3)$ significa que ha salido cara en ambas tiradas y el dado ha marcado el 3. Expresar el conjunto de todos los resultados posibles como producto cartesiano y determinar el número de resultados posibles.
5. ¿De cuántas maneras puede repartirse una baraja de bridge de 52 naipes en cuatro manos de 13 naipes? Exponer el razonamiento.
6. Se lanzan dos dados, uno rojo y uno blanco. Representar el resultado como un par ordenado (a, b), donde a representa el número de puntos del dado rojo, y b los del blanco. ¿Cuál es el número posible de pares ordenados (a, b)? ¿Cuántos de esos cumplen la condición de que la suma $a + b$ es:
   a) ¿par? b) ¿divisible por 3? c) ¿par o divisible por 3?
7. Una mano de poker contiene cinco naipes repartidos de una baraja de 52. ¿Cuántas manos distintas pueden repartirse que contengan:
   a) ¿dos parejas (por ejemplo, 2 reyes, 2 ases y un 3)?
   b) ¿un flux (cinco cartas de un mismo palo)?
   c) ¿una escalera de un mismo palo (cinco cartas consecutivas de un mismo palo, no incluyendo el «10», la sota, la reina, el rey y el as)?
   d) ¿una escalera real («10», sota, reina, rey, as de un mismo palo)?
8. En relación con el ejercicio 7. Calcular la probabilidad para que una mano de poker sea:
   a) un flux.
   b) una escalera de un mismo palo.
   c) una escalera real.
9. ¿Cuántos comités de 50 senadores norteamericanos pueden formarse que contengan:
   a) ¿un solo senador de Alaska?
   b) ¿dos senadores de Alaska?
10. Un comité de 50 senadores se elige al azar. Calcular la probabilidad de que queden incluidos los dos senadores por Alaska.
11. Se forman grupos de cuatro símbolos puestos en línea. Cada símbolo puede ser punto o raya. ¿Cuántos grupos pueden formarse?
12. ¿Cuántas palabras de k letras pueden formarse con un alfabeto que contiene n letras?
13. Demostrar que:

$$\text{a)}\quad \binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \cdots = 2^{n-1}.$$

b) $\binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}$.

14. Supongamos que un conjunto de pares ordenados (a, b) se ha construido eligiendo el primer componente en un conjunto de k elementos, por ejemplo, del conjunto $\{a_1, \ldots, a_k\}$, y el segundo componente en un conjunto de m elementos $\{b_1, \ldots, b_m\}$. Existen m pares con el primer componente a_1, esto es, $(a_1, b_1), \ldots, (a_1, b_m)$. Del mismo modo, hay m pares $(a_i, b_1), \ldots, (a_i, b_m)$ con el primer componente a_i. Por lo tanto, el número total de pares ordenados (a, b) es $m + m + \ldots + m$ (k sumandos). Esta suma vale km, que prueba la regla del cálculo secuencial para conjuntos de pares ordenados. Por inducción demostrar la regla para conjuntos de n-plas ordenadas.

13.12 Probabilidades condicionadas

Se lanza un dado correcto y se sabe que el resultado es un número par. ¿Cuál es la probabilidad de que este número sea divisible por 3? ¿Cuál es la probabilidad de que una criatura padezca daltonismo, en el supuesto de que es un niño? Estas preguntas pueden formularse en la siguiente forma: Sean A y B sucesos de un espacio muestral S. Si ocurre B, ¿cuál es la probabilidad de que ocurra A? Esto no es lo mismo que preguntar cuál es la probabilidad del suceso $A \cap B$. En realidad cuando $A = B$ la pregunta es: Si A ocurre, ¿cuál es la probabilidad de que A ocurra? La contestación sería en este caso 1, y ésta puede ser o no la probabilidad del suceso $A \cap B$. Para tratar tales problemas en general, volvamos al caso del dado.

Cuando nos ocupamos de cuestiones de probabilidad relativas al lanzamiento de un dado, utilizamos ordinariamente como espacio muestral $S = \{1, 2, 3, 4, 5, 6\}$ y asignamos la probabilidad $1/6$ a cada elemento de S. El suceso «divisible por 3» es el subconjunto $A = \{3, 6\}$ y el suceso «par» es el subconjunto $B = \{2, 4, 6\}$. Deseamos conocer la probabilidad de que un elemento esté en A, sabiendo ya que está en B. Ya que estamos interesados en resultados en los que el número es par, prescindimos de los resultados 1, 3, 5 y utilizamos, en lugar de S, el conjunto $B = \{2, 4, 6\}$ como espacio muestral. El suceso que nos interesa es tan sólo el conjunto de un elemento $\{6\}$, que es el único resultado del nuevo espacio muestral divisible por 3. Si todos los resultados de B se consideran igualmente probables, hay que asignar a cada uno de ellos la probabilidad $1/3$, luego la probabilidad de $\{6\}$ es también $1/3$.

Obsérvese que se ha resuelto el problema anterior empleando una idea muy elemental. Simplemente hemos cambiado el espacio muestral S por el B y se ha procedido a nueva asignación de probabilidades. Esto nos sugiere la manera de generalizar el procedimiento.

Sea $(S, \mathscr{B}, P)$ un espacio de probabilidad dado. Sean A y B dos sucesos y planteémonos la siguiente cuestión: «Si B acontece, ¿cuál es la probabilidad de que ocurra A?» Como en el anterior ejemplo, podemos cambiar el espacio mues-

tral S por el B y asignar nuevas probabilidades. Para el mismo B asignamos la probabilidad 1. Puesto que nos interesan los elementos de A que pertenecen al nuevo espacio muestral B, nuestro problema es calcular la probabilidad del suceso $A \cap B$ según las nuevas probabilidades asignadas. Esto es, si P' representa la función de probabilidad asociada al nuevo espacio muestral B, tenemos que calcular $P'(A \cap B)$.

Demostraremos ahora que si $P(B) \neq 0$ siempre podemos definir una función de probabilidad P' y un álgebra booleana $\mathscr{B}'$ de subconjuntos de B de modo que $(B, \mathscr{B}', P')$ sea un espacio de probabilidad. Como álgebra $\mathscr{B}'$ tomamos la colección de todos los conjuntos $T \cap B$ donde T es un conjunto que pertenece al álgebra booleana original $\mathscr{B}$. Es fácil comprobar que la $\mathscr{B}'$, así definida, es realmente un álgebra booleana. Una manera de definir una función de probabilidad P' sobre $\mathscr{B}'$ consiste en dividir cada una de las antiguas probabilidades por $P(B)$. Esto es, si $C \in \mathscr{B}'$ ponemos

$$P'(C) = \frac{P(C)}{P(B)}.$$

(Esto en la hipótesis de que $P(B) \neq 0$.) Lo que hemos hecho sólo ha sido modificar la escala, con todas las probabilidades multiplicadas por el factor $1/P(B)$. Es fácil comprobar que esta definición de P' nos da una medida de probabilidad. Es evidentemente no negativa y asigna la probabilidad 1 a B. La propiedad aditiva resulta inmediatamente de la de P.

Puesto que cada C de $\mathscr{B}'$ es de la forma $A \cap B$, donde A es un suceso en el espacio muestral original S, podemos dar a la definición de P' esta otra forma:

$$P'(A \cap B) = \frac{P(A \cap B)}{P(B)}.$$

Esta discusión sugiere que el cociente $P(A \cap B)/P(B)$ nos proporciona una justa medida de la probabilidad de que ocurra A, en el supuesto de que B haya ocurrido. La definición que sigue está hecha con esta intención:

DEFINICIÓN DE PROBABILIDAD CONDICIONADA. *Sea $(S, \mathscr{B}, P)$ un espacio de probabilidad y B un elemento tal que $P(B) \neq 0$. La probabilidad condicionada de que un suceso A ocurra, en el supuesto de que B ha ocurrido, se representa mediante el símbolo $P(A|B)$ (léase: «la probabilidad de A, dado B») y se define por la igualdad*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

La probabilidad condicionada $P(A|B)$ no está definida si $P(B) = 0$.

Los ejemplos que siguen ilustran el uso del concepto de probabilidad condicionada.

EJEMPLO 1. Consideremos una vez más el problema antes mencionado: Se lanza un dado y se sabe que el resultado es un número par. ¿Cuál es la probabilidad de que ese número sea divisible por 3? Podemos tomar por espacio muestral $S = \{1, 2, 3, 4, 5, 6\}$ y asignar la probabilidad $\frac{1}{6}$ a cada elemento de S. El suceso «par» es el conjunto $B = \{2, 4, 6\}$ y el suceso «divisible por 3» es el conjunto $A = \{3, 6\}$. Por consiguiente:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{3/6} = \frac{1}{3}.$$

Esto está de acuerdo con la anterior solución en la que se tomó B como espacio muestral y se asignó la probabilidad $\frac{1}{3}$ a cada elemento de B.

EJEMPLO 2. Se trata de un ejemplo típico usado en el Departamento de Biología de Caltech para prevenir contra los errores de las estadísticas superficiales. Para «demostrar» estadísticamente que la población de los Estados Unidos contiene más niños que niñas se pide a cada estudiante que cite el número de niños y niñas de su familia. Invariablemente, el número total de niños excede al de niñas. Las estadísticas en este caso son parciales ya que todos los estudiantes de bachillerato en Caltech son varones. Por lo tanto, la cuestión que aquí se considera no es precisamente la probabilidad de que una criatura sea niño, sino más bien la probabilidad de que una criatura sea niño en el supuesto que procede de una familia que por lo menos tiene un niño.

Para calcular las probabilidades en un ejemplo de este tipo consideremos una muestra de $4n$ familias, cada una con dos hijos. Supongamos que n familias tienen 2 niños, $2n$ familias un niño y una niña, y n familias 2 niñas. El espacio muestral S es el conjunto de las $8n$ criaturas de esas familias y asignemos la probabilidad $P(x) = 1/(8n)$ a cada x de S. Designemos por A el suceso «la criatura es niño» y por B «la criatura procede de una familia con un niño por lo menos». La probabilidad $P(A)$ es $1/2$. Análogamente $P(B) = 3/4$ ya que $3n$ de las $4n$ familias tienen por lo menos un niño. Por lo tanto la probabilidad de que una criatura sea niño, dado que procede de una familia con un niño por lo menos, es la probabilidad condicionada

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{1/2}{3/4} = \frac{2}{3}.$$

13.13 Independencia

Un concepto importante relacionado con el de probabilidad condicionada, es el de *independencia de sucesos,* que puede definirse como sigue:

DEFINICIÓN DE INDEPENDENCIA. *Dos sucesos A y B se llaman independientes (o estocásticamente independientes) si, y sólo si,*

(13.9) $$P(A \cap B) = P(A)P(B).$$

Si A y B son independientes, $P(A|B) = P(A)$ si $P(B) \neq 0$. Esto es, la probabilidad condicionada de A, dado B, es la misma que la probabilidad «absoluta» de A. Esta relación pone de manifiesto el significado de la independencia. El conocimiento de que B ha ocurrido, no influye en la probabilidad de que A ocurra.

EJEMPLO 1. Una carta es extraída de una baraja de 52. Cada carta tiene la misma probabilidad de ser seleccionada. Demostrar que los dos sucesos «sacar un as» y «sacar un corazón» son independientes.

Solución. Elegimos un espacio muestral S de 52 elementos y asignamos la probabilidad $\frac{1}{52}$ a cada elemento. El suceso A, «sacar un as» tiene la probabilidad $P(A) = \frac{4}{52} = \frac{1}{13}$. El suceso B, «sacar un corazón» tiene la probabilidad $P(B) = \frac{13}{52} = \frac{1}{4}$. El suceso $A \cap B$ significa «sacar el as de corazones», que tiene la probabilidad $\frac{1}{52}$, satisface la igualdad (13.9) y los sucesos A y B son independientes.

EJEMPLO 2. Se lanzan independientemente tres dados «correctos», de modo que cada combinación sea igualmente probable. Sea A el suceso consistente en que la suma de los dígitos obtenidos sea seis y B el que los tres dígitos sean distintos. Determinar si estos dos sucesos son independientes o no.

Solución. Como espacio muestral S tomamos el conjunto de todas las ternas (a, b, c) pudiendo a, b y c tomar los valores 1, 2, 3, 4, 5, 6. En S existen 6^3 elementos, y puesto que son igualmente probables asignamos la probabilidad $1/6^3$ a cada uno. El suceso A es el conjunto de todas las ternas (a, b, c) tales que $a + b + c = 6$. Por enumeración directa vemos que hay 10 ternas de esta índole, a saber:

$$(1, 2, 3), (1, 3, 2), (1, 1, 4), (1, 4, 1),$$

$$(2, 1, 3), (2, 3, 1), (2, 2, 2),$$

$$(3, 1, 2), (3, 2, 1),$$

$$(4, 1, 1).$$

El suceso B consta de todas las ternas (a, b, c) para las que $a \neq b$, $b \neq c$, y $a \neq c$. Existen $6 \cdot 5 \cdot 4 = 120$ elementos en B. Exactamente seis de esos elementos pertenecen a A, de manera que $A \cap B$ tiene seis elementos. Por lo tanto,

$$P(A \cap B) = \frac{6}{6^3}, \qquad P(A) = \frac{10}{6^3}, \qquad \text{y} \qquad P(B) = \frac{120}{6^3}.$$

En este caso $P(A \cap B) \neq P(A)P(B)$; luego los sucesos A y B no son independientes.

Para más de dos elementos se define la independencia del modo siguiente. Una colección finita $\mathscr{A}$ de n sucesos se dice que es independiente si los sucesos satisfacen la propiedad multiplicativa

$$(13.10) \qquad P\left(\bigcap_{k=1}^{m} A_k\right) = \prod_{k=1}^{m} P(A_k)$$

para *toda* subcolección finita $\{A_1, A_2, \ldots, A_m\}$, donde m puede tomar los valores $m = 2, 3, \ldots, n$, y los conjuntos A_i pertenecen a $\mathscr{A}$.

Cuando $\mathscr{A}$ consta exactamente de tres sucesos A, B y C, la condición de independencia (13.10) exige que

$$(13.11) \quad P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C), \quad P(B \cap C) = P(B)P(C),$$

y

$$(13.12) \qquad P(A \cap B \cap C) = P(A)P(B)P(C).$$

Puede pensarse que las tres igualdades (13.11) son suficientes para que se satisfaga (13.12) o, en otras palabras, que la independencia de tres sucesos es una consecuencia de la independencia *a pares*. Esto no es cierto, como puede verse con el siguiente ejemplo:

Cuatro tickets con las marcas a, b, c y abc, se colocan en una caja. Se extrae un ticket al azar, y el espacio muestral es:

$$S = \{a, b, c, abc\}.$$

Los sucesos A, B y C se definen como sigue:

$$A = \{a, abc\}, \qquad B = \{b, abc\}, \qquad C = \{c, abc\}.$$

Esto significa, que el suceso X consiste en la extracción de un ticket que contiene

la letra x. Es fácil comprobar que cada una de las tres igualdades (13.11) se satisface de manera que los sucesos A, B, y C son independientes a pares. Sin embargo, (13.12) no se satisface y por consiguiente los *tres* sucesos no son independientes. Los cálculos son sencillos y los dejamos como ejercicio para el lector.

13.14 Ejercicios

1. Sean A y B dos sucesos con $P(A) \neq 0$, $P(B) \neq 0$. Demostrar que

$$P(A \cap B) = P(B)P(A \mid B) = P(A)P(B \mid A). \tag{13.13}$$

Algunas veces es más fácil calcular las probabilidades $P(A)$ y $P(B|A)$ directamente enumerando los casos que calcular $P(A \cap B)$. Cuando esto ocurre, la igualdad (13.13) nos da un método adecuado para calcular $P(A \cap B)$. En el próximo ejercicio damos un ejemplo.

2. Una urna contiene siete bolas blancas y tres negras. Una segunda urna contiene cinco blancas y cinco negras. Se extrae al azar una bola de la primera urna y se coloca en la segunda. A continuación se extrae al azar una bola de la segunda urna. Representemos con A el suceso «bola negra en la primera extracción» y con B el suceso «bola negra en la segunda extracción».
   a) Calcular las probabilidades $P(A)$ y $P(B|A)$ directamente por enumeración de los casos posibles. Utilizar la igualdad (13.13) para calcular $P(A \cap B)$.
   b) Calcular directamente $P(A \cap B)$ enumerando todos los pares de extracciones posibles.
3. a) Sean A_1, A_2, A_3 tres sucesos tales que $P(A_1 \cap A_2) \neq 0$. Demostrar que

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2).$$

   b) Generalizar por inducción este resultado como sigue: Si $A_1, A_2, \ldots, A_n$ son n sucesos $(n \geq 2)$ tales que $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \neq 0$, entonces,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

4. Un comité de 50 senadores se elige al azar. Hallar la probabilidad de que los dos senadores por Alaska queden incluidos, dado que por lo menos uno ya lo está.
5. Una urna contiene cinco bolas doradas y siete azules. Se extraen al azar dos bolas (sin devolverlas a la urna). Si la primera bola es dorada, calcular la probabilidad de que la segunda sea también dorada.
6. Un baraja es repartida en cuatro manos de 13 naipes cada una. Si una mano tiene exactamente siete espadas, ¿cuál es la probabilidad de que una determinada de las otras manos contenga a) por lo menos una espada, b) por lo menos dos espadas, c) un palo completo?
7. Demostrar que $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$.
8. Sean $A_1, A_2, \ldots, A_n$ n sucesos disjuntos cuya reunión es el espacio muestral S completo. Para todo suceso E tenemos la igualdad

$$E = E \cap S = E \cap \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} (E \cap A_i).$$

Esta igualdad establece que E puede ocurrir tan sólo en conjunción con algún A_i. Demostrar que

a) $P(E) = \sum_{i=1}^{n} P(E \cap A_i)$.

b) $P(E) = \sum_{i=1}^{n} P(E \mid A_i)P(A_i)$.

Esta fórmula es útil cuando las probabilidades condicionadas $P(E|A_i)$ son más sencillas de calcular que $P(E)$.

9. Una moneda «correcta» es lanzada repetidamente. Sale cara en los seis primeros lanzamientos. ¿Cuál es la probabilidad de que salga cara en el séptimo?
10. Dados los sucesos independientes A y B cuyas probabilidades no son 0 ni 1. Demostrar que A' y B' son independientes. ¿Ocurre lo mismo si A o B tienen probabilidad 0 ó 1?
11. Dados los sucesos independientes A y B. Probar o rechazar, según convenga en cada caso, que:
    a) A' y B son independientes.
    b) $A \cup B$ y $A \cap B$ son independientes.
    c) $P(A \cup B) = 1 - P(A')P(B')$.
12. Si $A_1, A_2, \ldots, A_n$ son sucesos independientes, demostrar que

$$P\left(\bigcup_{i=1}^{n} A_i\right) + \prod_{i=1}^{n} P(A_i') = 1 .$$

13. Si tres sucesos A, B, y C son independientes, demostrar que $A \cup B$ y C son independientes. [*Indicación*. Utilizar el resultado del ejercicio 7 para demostrar que $P(A \cup B \mid C) = P(A \cup B)$.]
14. Sean A y B dos sucesos, ninguno de los cuales tiene probabilidad 0. Demostrar que son ciertas o que no lo son cada una de las afirmaciones siguientes:
    a) Si A y B son disjuntos, A y B son independientes.
    b) Si A y B son independientes, A y B son disjuntos.
15. Se lanza un dado dos veces, el espacio muestral consta de 36 posibles pares de resultados (a, b) cada uno con probabilidad $\frac{1}{36}$. Sean A, B y C los sucesos siguientes:

$$A = \{(a, b) \mid a \text{ es impar}\}, \quad B = \{(a, b) \mid b \text{ es impar}\}, \quad C = \{(a, b) \mid a + b \text{ es impar}\}.$$

    a) Calcular $P(A), P(B), P(C), P(A \cap B), P(A \cap C), P(B \cap C)$, y $P(A \cap B \cap C)$.
    b) Demostrar que A, B y C son independientes a pares.
    c) Demostrar que A, B y C no son independientes.

13.15 Experimentos o pruebas compuestas

Volvamos ahora al problema de de Méré mencionado en la introducción —si es lo mismo apostar la misma cantidad a favor o en contra de la aparición por lo menos de un «doble seis» en 24 tiradas de un par de dados. Tratemos el problema de un modo más general: ¿Cuál es la probabilidad de conseguir un doble seis

por lo menos una vez en n tiradas de un par de dados? ¿Esta probabilidad es mayor o menor que un medio cuando $n = 24$?

Consideremos en primer lugar la prueba consistente en lanzar un par de dados una sola vez. El resultado de este juego puede representarse mediante pares ordenados (a, b) en los que a y b recorren los valores 1, 2, 3, 4, 5, 6. El espacio muestral S consta de 36 de esos pares. Puesto que los dados son correctos asignamos a cada par de S la probabilidad $1/36$.

Supongamos que lanzamos los dados n veces. La sucesión de las n pruebas es una prueba compuesta que queremos describir matemáticamente. Por ello necesitamos un nuevo espacio muestral y una correspondiente medida de probabilidad. Consideremos los resultados del nuevo juego como n-plas ordenadas $(x_1, \ldots, x_n)$, en donde cada componente x_i es uno de los resultados del espacio muestral original S. Es decir, el espacio muestral para la prueba compuesta es el producto cartesiano $S \times \ldots \times S$, que se representa por S^n. El conjunto S^n tiene 36^n elementos, y asignamos igual probabilidad a cada elemento:

$$P(x) = \frac{1}{36^n} \qquad \text{si} \quad x \in S^n.$$

Nos interesa el suceso «por lo menos un doble seis en n tiradas.» Designémoslo por A. En este caso es más sencillo calcular la probabilidad del suceso complementario A', que significa «ningún doble seis en n tiradas.» Cada elemento de A' es una n-pla, cada uno de cuyos componentes pueden ser cualquier elemento de S excepto (6, 6). Por consiguiente existen 35 valores posibles para cada componente y por tanto $(35)^n$ n-plas en total en A'. Puesto que cada elemento de A' tiene probabilidad $(\frac{1}{36})^n$, la suma de todas las probabilidades puntuales en A' es $(\frac{35}{36})^n$. Esto nos da

$$P(A) = 1 - P(A') = 1 - (\tfrac{35}{36})^n.$$

Para contestar a la pregunta de de Méré tenemos que decidir si $P(A)$ es mayor o menor que $1/2$ cuando $n = 24$. La desigualdad $P(A) \geqslant 1/2$ es equivalente a $1 - (\frac{35}{36})^n \geqslant 1/2$, o $(\frac{35}{36})^n \leqslant 1/2$. Tomando logaritmos encontramos

$$n \log 35 - n \log 36 \leq -\log 2, \qquad \text{o} \qquad n \geq \frac{\log 2}{\log 36 - \log 35} = 24.6+ .$$

Por consiguiente $P(A) < 1/2$ cuando $n = 24$ y $P(A) > 1/2$ cuando $n \geqslant 25$. *No es* ventajosa una apuesta de una cantidad al suceso de que por lo menos se presente un doble seis en 24 tiradas, frente a la apuesta de la misma cantidad al suceso contrario.

Esta discusión sugiere un método general para tratar los experimentos sucesivos. Si una prueba se repite dos o más veces, el resultado puede considerarse

como una prueba compuesta. Más general, una prueba compuesta puede ser el resultado de ejecutar dos o más pruebas distintas sucesivamente. Cada una de las pruebas individuales puede estar relacionada con cada una de las otras o pueden ser estocásticamente independientes, en el sentido de que la probabilidad del resultado de cada una de ellas no depende de los resultados de las otras.

Por simplicidad, discutiremos cómo se pueden combinar *dos* pruebas independientes en una prueba compuesta. La generalización a más de dos experiencias será evidente.

Para asociar el espacio de probabilidad natural a una prueba o experiencia compuesta, debemos definir el nuevo espacio muestral S, la correspondiente álgebra booleana $\mathscr{B}$ de subconjuntos de S, y la medida de probabilidad P sobre $\mathscr{B}$. Como en el ejemplo anterior, usamos el concepto de *producto cartesiano.*

Sean $(S_1, \mathscr{B}_1, P_1)$ y $(S_2, \mathscr{B}_2, P_2)$. dos espacios de probabilidad asociados a dos experiencias E_1 y E_2. Con E representamos la experiencia o prueba compuesta para las que el espacio muestral S es el producto cartesiano $S_1 \times S_2$. Un resultado de E es el par (x, y) de S, donde el primer componente x es un resultado de E_1 y el segundo y un resultado de E_2. Si S_1 tiene n elementos y S_2 m, el producto $S_1 \times S_2$ tendrá nm elementos.

Como nueva álgebra booleana $\mathscr{B}$ tomamos la colección de todos los subconjuntos de S. A continuación definimos la función de probabilidad P. Ya que S es finito podemos definir $P(x, y)$ para cada punto (x, y) de S y utilizar la aditividad al definir P para los subconjuntos de S. Las probabilidades $P(x, y)$ pueden asignarse de varias maneras. Sin embargo, si dos pruebas E_1 y E_2 son estocásticamente *independientes,* definimos P mediante la ecuación

(13.14) $$P(x, y) = P_1(x)P_2(y) \quad \text{para cada } (x, y) \text{ de } S.$$

Se justifica esta definición del siguiente modo. Consideremos dos sucesos particulares A y B del nuevo espacio S,

$$A = \{(x_1, y_1), (x_1, y_2), \ldots, (x_1, y_m)\}$$

y

$$B = \{(x_1, y_1), (x_2, y_1), \ldots, (x_n, y_1)\}.$$

Esto es, A es el conjunto de todos los pares de $S_1 \times S_2$ cuyo primer elemento es x_1, y B es el conjunto de todos los pares cuyo segundo elemento es y_1. La intersección de los dos conjuntos A y B es el conjunto de un solo elemento $\{(x_1, y_1)\}$. Si presentimos que el primer resultado x_1 no debe influir en el resultado y_1 parece razonable exigir que los sucesos A y B sean independientes. Esto significa que habrá que definir la nueva función de probabilidad P de manera que

(13.15) $$P(A \cap B) = P(A)P(B).$$

Si decidimos la forma de asignar las probabilidades $P(A)$ y $P(B)$, la ecuación (13.15) nos dirá cómo asignar la probabilidad $P(A \cap B)$, esto es, la probabilidad $P(x_1, y_1)$. Se presenta el suceso A si y sólo si el resultado de la primera prueba es x_1. Puesto que $P_1(x_1)$ es su probabilidad, parece natural asignar el valor $P_1(x_1)$ también a $P(A)$. Análogamente, asignamos a $P(B)$ el valor $P_2(y_1)$. La ecuación (13.15) nos da entonces

$$P(x_1, y_1) = P_1(x_1)P_2(y_1).$$

Todo esto es, naturalmente, tan solo una justificación para la asignación de probabilidades (13.14). El único camino para decidir si (13.14) es o no una aceptable asignación de probabilidades puntuales es ver si se cumplen las propiedades fundamentales de las medidas de probabilidad. Cada número $P(x, y)$ es no negativo, y la suma de todas las probabilidades puntuales es igual a 1, puesto que tenemos

$$\sum_{(x,y)\in S} P(x, y) = \sum_{x\in S_1} P_1(x) \cdot \sum_{y\in S_2} P_2(y) = 1 \cdot 1 = 1 .$$

Cuando decimos que una prueba compuesta E está determinada por dos pruebas E_1 y E_2 estocásticamente independientes, queremos decir que el espacio de probabilidad $(S, \mathscr{B}, P)$ está definido como acabamos de explicar, tal «independencia» queda reflejada en el hecho de que $P(x, y)$ es igual al producto $P_1(x)$ $P_2(y)$. Puede demostrarse que la asignación de probabilidades (13.14) implica la igualdad

$$P(U \times V) = P_1(U)P_2(V) \tag{13.16}$$

para todo par de subconjuntos U de $\mathscr{B}_1$ y V de $\mathscr{B}_2$. (Véase el ejercicio 12 de la sección 13.23 donde se esboza la demostración.) De esta forma deduciremos algunas consecuencias importantes.

Sea A un suceso (de la prueba compuesta E) de la forma

$$A = C_1 \times S_2,$$

donde $C_1 \in \mathscr{B}_1$. Cada resultado de A es un par ordenado (x, y) siendo x un resultado de C_1 (en la primera prueba E_1) mientras que y puede ser cualquier resultado de S_2 (en la segunda prueba E_2). Si aplicamos (13.16) encontramos:

$$P(A) = P(C_1 \times S_2) = P_1(C_1)P_2(S_2) = P_1(C_1),$$

ya que $P_2(S_2) = 1$. De este modo la definición de P asigna la misma probabilidad a A que la asignada por P_1 a C_1. Por esta razón, se dice que un tal suceso A está

determinado mediante la primera prueba E_1. Análogamente, si B es un suceso de E de la forma

$$B = S_1 \times C_2,$$

perteneciendo $C_2 \in \mathscr{B}_2$, tenemos:

$$P(B) = P(S_1 \times C_2) = P_1(S_1)P_2(C_2) = P_2(C_2)$$

y se dice que B está *determinado por la segunda prueba* E_2. Demostraremos ahora, utilizando (13.16), que tales sucesos A y B son *independientes*. Esto es, tenemos:

$$(13.17) \qquad P(A \cap B) = P(A)P(B).$$

Ante todo notemos que

$$\begin{aligned} A \cap B &= \{(x, y) \mid (x, y) \in C_1 \times S_2 \ \text{ y } \ (x, y) \in S_1 \times C_2\} \\ &= \{(x, y) \mid x \in C_1 \text{ y } \ y \in C_2\} \\ &= C_1 \times C_2 . \end{aligned}$$

Luego, en virtud de (13.16), tenemos:

$$(13.18) \qquad P(A \cap B) = P(C_1 \times C_2) = P_1(C_1)P_2(C_2).$$

Puesto que $P_1(C_1) = P(A)$ y $P_2(C_2) = P(B)$ obtenemos (13.17). Obsérvese que la igualdad (13.18) también demuestra que podemos calcular la probabilidad $P(A \cap B)$ como producto de las probabilidades en cada uno de los espacios muestrales S_1 y S_2; luego no son precisos cálculos con probabilidades en la prueba compuesta.

La generalización a pruebas o experiencias compuestas de n pruebas E_1, E_2, $\ldots, E_n$ se deduce en la misma forma. Los puntos en el nuevo espacio muestral son n-plas $(x_1, x_2, \ldots, x_n)$ y las probabilidades se definen como producto de las probabilidades de los resultados particulares

$$(13.19) \qquad P(x_1, x_2, \ldots, x_n) = P_1(x_1)P_2(x_2) \cdots P_n(x_n).$$

Cuando se adopta esta definición de P decimos que E está determinado por *n pruebas independientes* E_1, $E_2, \ldots, E_n$. En el caso particular en el que todas las pruebas están asociadas al mismo espacio de probabilidad, la prueba compuesta E es un ejemplo de *pruebas independientes repetidas bajo idénticas condiciones*. En la sección siguiente se considera un ejemplo.

13.16 Pruebas de Bernoulli

Un ejemplo importante de prueba compuesta lo estudió Jacobo Bernoulli y lo conocemos por el nombre de *sucesión de pruebas de Bernoulli.* Se trata de una sucesión de pruebas repetidas ejecutadas en las mismas condiciones, siendo cada resultado estocásticamente independiente de las demás. Cada prueba tiene exactamente dos resultados posibles, corrientemente llamados «éxito» y «fallo»; la probabilidad del éxito se representa por p y la del fallo con q. Naturalmente, $q = 1-p$. El teorema principal relacionado con las sucesiones de Bernoulli es el siguiente:

TEOREMA 13.3. FÓRMULA DE BERNOULLI. *La probabilidad de k éxitos en n pruebas de Bernoulli es*

$$\binom{n}{k} p^k q^{n-k}, \tag{13.20}$$

donde $\binom{n}{k}$ *representa el coeficiente binomial,* $\binom{n}{k} = \dfrac{n!}{k!\,(n-k)!}$.

Demostración. Representemos el «éxito» con S y el «fallo» con F y consideremos una sucesión particular de n resultados. Esto puede representarse mediante una n-pla.

$$(x_1, x_2, \ldots, x_n),$$

donde cada x_i es S o F. El suceso A en el que estamos interesados es la colección de todas las n-plas que contienen exactamente k veces la S y $n-k$ veces la F. Calculemos la probabilidad de una n-pla determinada de A. La probabilidad de cada S es p, y la de cada F es q. Luego, en virtud de (13.19), la probabilidad de cada n-pla de A es el producto de k factores iguales a p por $n-k$ factores q. Esto es,

$$P(x_1, x_2, \ldots, x_n) = p^k q^{n-k} \quad \text{si} \quad (x_1, x_2, \ldots, x_n) \in A.$$

Por lo tanto, para calcular $P(A)$ tan sólo tenemos que contar el número de elementos de A y multiplicar dicho número por $p^k q^{n-k}$. Pero el número de elementos de A es simplemente el número de maneras de colocar k veces la S en las n posiciones posibles de *n-pla.* Esto es lo mismo que el número de subconjuntos de k elementos que pueden formarse con un conjunto de n elementos; sabemos ya que este número es $\binom{n}{k}$. De ahí que, si sumamos las probabilidades correspondientes

a todos los puntos de A obtenemos:

$$P(A) = \binom{n}{k} p^k q^{n-k}.$$

EJEMPLO 1. Se lanza 50 veces una moneda. Calcular la probabilidad de que salgan 25 caras exactamente.

Solución. Interpretamos este juego como una sucesión de 50 pruebas de Bernoulli, en las que «éxito» significa «cara» y «fallo» será «cruz». Supuesta la moneda correcta asignamos las probabilidades $p = q = 1/2$, y la fórmula (13.20) nos da $\binom{50}{k}(\frac{1}{2})^{50}$ como probabilidad de obtener exactamente k veces cara en 50 tiradas. En particular, si $k = 25$ obtenemos:

$$\binom{50}{25}\left(\frac{1}{2}\right)^{50} = \frac{50!}{25!\,25!}\left(\frac{1}{2}\right)^{50}.$$

Para expresar este número en forma decimal es mejor utilizar logaritmos, y manejando tablas de logaritmos de factoriales. Designando el número buscado por P, en logaritmos base 10 obtenemos

$$\begin{aligned}\log P &= \log 50! - 2\log 25! - 50\log 2\\ &= 64.483 - 50.381 - 15.052 = -0.950 = 0.05 - 1.00\\ &= \log 1.12 - \log 10 = \log 0.112,\end{aligned}$$

así que $P = 0{,}112$.

EJEMPLO 2. ¿Cuál es la probabilidad de conseguir r éxitos por lo menos en n pruebas de Bernoulli?

Solución. Sea A_k el suceso «obtener exactamente k éxitos en n pruebas». El suceso E que analizamos es la reunión

$$E = A_r \cup A_{r+1} \cup \cdots \cup A_n.$$

Ya que los A_k son disjuntos, encontramos:

$$P(E) = \sum_{k=r}^{n} P(A_k) = \sum_{k=r}^{n} \binom{n}{k} p^k q^{n-k}.$$

Puesto que,

$$\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k} = (p+q)^n = 1\,,$$

la probabilidad del suceso complementario E' puede calcularse así:

$$P(E') = 1 - P(E) = \sum_{k=0}^{r-1}\binom{n}{k}p^k q^{n-k}.$$

Esta última suma nos da la probabilidad de lograr por lo menos $r - 1$ éxitos en n pruebas.

13.17 Número más probable de éxitos en *n* pruebas de Bernoulli

Un par de dados correctos es lanzado 28 veces. ¿Cuál es el número más probable de sietes? Para resolver este problema designemos con $f(k)$ la probabilidad de obtener exactamente k sietes en 28 tiradas. La probabilidad de conseguir un siete en una tirada es $1/6$. La fórmula de Bernoulli nos dice que

$$f(k) = \binom{28}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{28-k}.$$

Queremos determinar qué valor (o valores) de k entre los valores $k = 0, 1, 2, \ldots, 28$ hacen máximo a $f(k)$. El siguiente teorema resuelve la cuestión para cualquier sucesión de pruebas de Bernoulli.

TEOREMA 13.4. *Dados un entero $n \geq 1$ y un número real p, $0 < p < 1$, consideremos el conjunto de números*

$$f(k) = \binom{n}{k}p^k(1-p)^{n-k}, \quad \textit{para} \quad k = 0, 1, \ldots, n\,.$$

a) *Si $(n+1)p$ no es entero, el máximo de $f(k)$ se presenta exactamente para un valor de k:*

$$k = [(n+1)p], \qquad \textit{el mayor entero} < (n+1)p.$$

b) *Si $(n+1)p$ es entero, el máximo de $f(k)$ se presenta exactamente para dos valores de k:*

$$k = (n+1)p \qquad \textit{y} \qquad k = (n+1)p - 1\,.$$

Demostración. Para estudiar el comportamiento de $f(k)$ consideremos la razón

$$r(k) = \frac{f(k)}{f(k+1)} = \frac{k+1}{n-k}\frac{1-p}{p}$$

para $k = 0,1,\ldots,n-1$. La función $r(k)$ es creciente en sentido estricto por lo cual tenemos

$$0 < r(0) < r(1) < \cdots < r(n-1).$$

Vamos a considerar seis casos, representados en la figura 13.2. En los tres primeros demostramos que $f(k)$ toma su valor máximo para un solo valor de k. En los restantes $f(k)$ alcanzan su máximo para dos valores consecutivos de k.

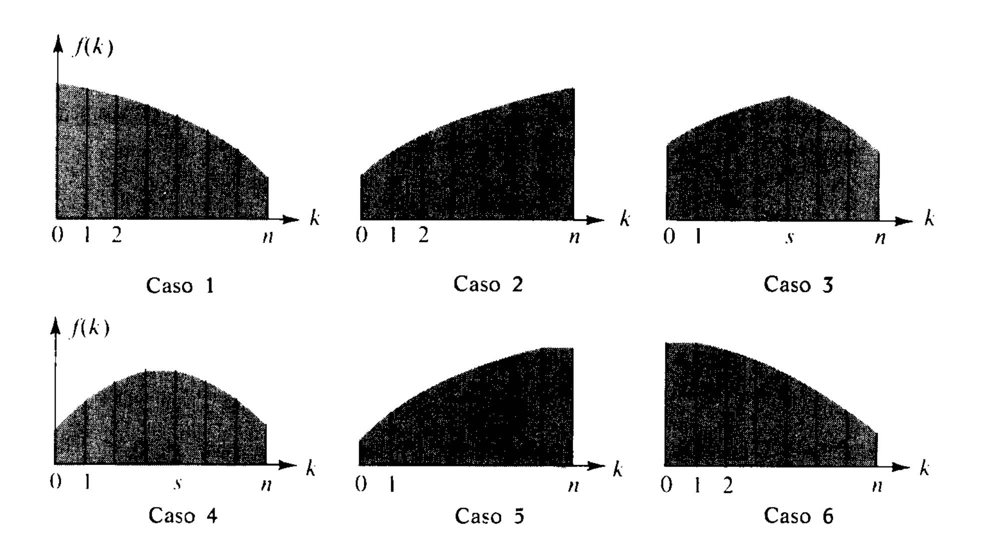


FIGURA 13.2 *Cálculo del número más probable de éxitos en n pruebas de Bernoulli.*

CASO 1. $r(0) > 1$. En este caso $r(k) > 1$ para todo k con lo que

$$f(0) > f(1) > \cdots > f(n).$$

Por consiguiente el valor máximo de $f(k)$ se presenta sólo para $k = 0$. También, $r(0) = (1-p)/(np) > 1$, así que $1 - p > np$, $(n+1)p < 1$, luego $[(n+1)p] = 0$.

CASO 2. $r(n-1) < 1$. En este caso $r(k) < 1$ para todo k con lo que $f(0) < f(1) < \ldots < f(n)$ y el valor máximo de $f(k)$ se presenta sólo para $k = n$. Puesto que $r(n-1) = n(1-p)/p < 1$, tenemos $n - np < p$, luego $n < (n+1)p < n+1$, así que $[(n+1)p] = n$.

CASO 3. $r(0) < 1$, $r(n-1) > 1$, y $r(k) \neq 1$ *para todo* k. En este caso hay un único entero s, $0 < s < n$, tal que $r(s-1) < 1$ y $r(s) > 1$. La función $f(k)$ es creciente en el intervalo $0 \leq k \leq s$ y decreciente en el intervalo $s \leq k \leq n$. Por consiguiente $f(k)$ tiene un solo máximo en $k = s$. Puesto que $r(s-1) = s(1-p)/(np - sp + p) < 1$ tenemos $s < (n+1)p$. La desigualdad $r(s) > 1$ demuestra que $(n+1)p < s+1$, luego $[(n+1)p] = s$.

Obsérvese que en cada uno de los tres primeros casos el valor máximo de $f(k)$ se presenta cuando $k = [(n+1)p]$; asimismo $(n+1)p$ no es entero en ninguno de esos casos.

CASO 4. $r(0) < 1$, $r(n-1) > 1$, y $r(s-1) = 1$ *para un cierto* s, $2 \leq s < n$. En este caso $f(k)$ crece para $0 \leq k \leq s-1$ y decrece para $s \leq k \leq n$. El valor máximo de $f(k)$ se presenta dos veces, cuando $k = s-1$ y cuando $k = s$. La ecuación $r(s-1) = 1$ implica $(n+1)p = s$.

CASO 5. $r(n-1) = 1$. En este caso $r(k) < 1$ para $k \leq n-2$, así que $f(k)$ crece en el intervalo $0 \leq k \leq n-1$, y $f(n-1) = f(n)$. Luego $f(k)$ tiene dos máximos, cuando $k = n-1$ y cuando $k = n$. La ecuación $r(n-1) = 1$ implica $(n+1)p = n$.

CASO 6. $r(0) = 1$. En este caso $r(k) > 1$ para $k \geqslant 1$, así que $f(k)$ decrece en el intervalo $1 \leq k \leq n$. Los máximos de $f(k)$ son dos, cuando $k = 0$ y cuando $k = 1$. La ecuación $r(0) = 1$ implica $(n+1)p = 1$.

En cada uno de los tres últimos casos el valor máximo de $f(k)$ ocurre para $k = (n+1)p$ y para $k = (n+1)p - 1$. Esto completa la demostración.

EJEMPLO 1. Un par de dados se lanza 28 veces. ¿Cuál es el número más probable de sietes?

Solución. Se aplica el teorema 13.4 con $n = 28$, $p = \frac{1}{6}$, y $(n+1)p = \frac{29}{6}$. Éste no es un entero con lo cual el valor máximo de $f(k)$ se presenta para $k = [\frac{29}{6}] = 4$.

Observación: Si se echan los dados 29 veces hay dos soluciones, $k=4$ y $k=5$

EJEMPLO 2. Hallar el menor n tal que si se tiran dos dados correctos n veces la probabilidad de conseguir cuatro sietes es por lo menos tan grande como la de obtener cualquier otro número de sietes.

Solución. Tomamos $p = 1/6$ en el teorema 13.4. Queremos que el máximo de $f(k)$ ocurra cuando $k=4$. Esto exige o bien que $[(n+1)p]=4$, $(n+1)p=4$, o $(n+1)p-1=4$. El menor valor de n que satisface cualquiera de esas relaciones es $n = 23$.

13.18 Ejercicios

1. Se lanza una moneda dos veces, la probabilidad de cara en la primera tirada sea p_1 y en la segunda tirada p_2. Consideremos esto como una prueba compuesta determinada por dos pruebas estocásticamente independientes, y sea el espacio muestral

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

a) Calcular la probabilidad de cada elemento de S.
b) ¿Puede ser la asignación de las probabilidades p_1 y p_2 de modo que

$$P(H, H) = \tfrac{1}{9}, \qquad P(H, T) = P(T, H) = \tfrac{2}{9}, \qquad P(T, T) = \tfrac{4}{9}?$$

c) ¿Pueden ser asignadas p_1 y p_2 de modo que

$$P(H, H) = P(T, T) = \tfrac{1}{3}, \qquad P(H, T) = P(T, H) = \tfrac{1}{6}?$$

d) Considerar los cuatro sucesos siguientes (subconjuntos de S):
H_1: cara en la primera tirada,
H_2: cara en la segunda tirada,
T_1: cruz en la primera tirada,
T_2: cruz en la segunda tirada.
Determinar qué pares de esos cuatro sucesos son independientes.

En cada uno de los ejercicios del 2 al 12 determinar el espacio muestral, la asignación de probabilidades, y el suceso cuya probabilidad se calcule.

2. Un estudiante debe rendir un examen consistente en 10 preguntas. No está preparado para ello y se propone acertar las preguntas contestándolas al azar. Por ejemplo, puede lanzar una buena moneda y utilizar el resultado para determinar su pronóstico.
a) Calcular la probabilidad de que acierte correctamente por lo menos cinco veces.
b) Calcular la probabilidad de que acierte correctamente por lo menos nueve veces.
c) ¿Cuál es el menor valor de n tal que la probabilidad de acertar por lo menos n respuestas correctas es menor que $\frac{1}{2}$?
3. Diez dados correctos se lanzan juntos. ¿Cuál es la probabilidad de obtener exactamente tres seises?
4. Se lanza cinco veces una moneda correcta. ¿Cuál es la probabilidad de obtener a) tres caras exactamente, b) por lo menos tres caras, c) por lo menos una cara?
5. Un hombre afirmaba poseer una varilla con la que podía localizar yacimientos de petróleo. El Departamento de Geología de Caltech realizó el siguiente experimento para poner a prueba su afirmación. Fue colocado en una habitación donde había 10 barriles precintados. Se le advirtió que cinco de ellos contenían petróleo y los otros cinco agua.

Su trabajo consistió en decidir cuáles contenían petróleo y cuáles agua.

a) ¿Cuál es la probabilidad de que localizara correctamente los cinco barriles de petróleo tan sólo por azar?

b) ¿Cuál la de que localizara por lo menos tres de los barriles de petróleo únicamente por azar?

6. Una anciana de Pasadena afirma que probando una taza de té con leche puede decir qué fue lo primero que se puso en la taza: el té o la leche. Tal afirmación se pone a prueba haciéndole degustar y clasificar 10 pares de tazas de té, conteniendo cada par una taza de té servida de cada una de las dos maneras citadas. Sea p su probabilidad «cierta» de clasificar un par de tazas correctamente. (Si ella es hábil, p es substancialmente mayor que $\frac{1}{2}$; si no lo es, $p \leq \frac{1}{2}$.) Se supone que los 10 pares de tazas son clasificadas bajo condiciones idénticas e independientes.

 a) Calcular, en función de p, la probabilidad de que clasifique correctamente por lo menos ocho de los 10 pares de tazas.

 b) Valorar esta probabilidad explícitamente cuando $p = \frac{1}{2}$.

7. (Otro problema del caballero de Méré.) Determinar si es o no ventajoso jugar apostando cantidades iguales, a que por lo menos aparece un 6 en cuatro tiradas de un dado correcto. [*Indicación.* Probar que la probabilidad de sacar por lo menos un 6 en n tiradas es $1 - (\frac{5}{6})^n$.]

8. Una urna contiene a bolas blancas y b bolas negras. Si $k \leq n$, calcular la probabilidad de sacar k bolas blancas en n extracciones, si cada bola es devuelta a la urna antes de sacar la siguiente.

9. Se lanzan dos dados ocho veces. Calcular la probabilidad de que la suma sea 11 exactamente tres veces.

10. Se lanza una moneda 10 veces o 10 monedas una vez y se cuenta el número de caras. Encontrar la probabilidad de obtener por lo menos seis caras.

11. Después de una larga serie de tests aplicados a cierto tipo de cohete se ha determinado que aproximadamente en un 5 % de pruebas se producirá un mal funcionamiento que será la causa de que el cohete fracase. Calcular la probabilidad de que en 10 pruebas haya por lo menos un fallo.

12. Se lanza repetidamente una moneda. Calcular la probabilidad de que el número total de caras sea por lo menos 6 antes de que el número total de cruces sea 5.

13. El ejercicio 12 puede generalizarse como sigue: Demostrar que la probabilidad de que se produzcan por lo menos m éxitos antes que n fallos en una sucesión de pruebas de Bernoulli es:

$$\sum_{k=m}^{m-n-1} \binom{m+n-1}{k} p^k q^{m+n-k-1}.$$

14. Determinar todos los valores de n con la siguiente propiedad: Si se tira n veces un par de dados correctos, la probabilidad de obtener exactamente diez sietes es por lo menos tan grande como la de obtener cualquier otro número de sietes.

15. Una máquina tragamonedas binaria tiene tres ruedas idénticas e independientes. Cuando se juega con la máquina los resultados que pueden obtenerse son ternas ordenadas (x, y, z), en donde cada una de las letras x, y, z puede ser 0 ó 1. En cada rueda la probabilidad de 0 es p y la probabilidad de 1 es $1 - p$, en donde $0 < p < 1$. La máquina paga 2$ si el resultado es $(1, 1, 1)$ o $(0, 0, 0)$; paga 1$ si se obtiene el resultado $(1, 1, 0)$; en cualquier otro caso no paga nada. Designemos con $f(p)$ la probabilidad de que la máquina pague un dólar o más cuando se haga una tirada.

 a) Calcular $f(p)$.

b) Definamos el «pago total» como la suma $\sum_{x\in S} g(x)P(x)$, en donde S es el espacio muestral, $P(x)$ es la probabilidad del resultado x, y $g(x)$ el número de dólares pagados por el resultado x. Calcular el valor de p para el cual el «pago total» sea mínimo.

13.19 Conjuntos numerables y no numerables

Hasta aquí sólo hemos considerado el concepto de probabilidad para espacios muestrales finitos. Queremos ahora extender la teoría a espacios muestrales infinitos. Para ello es necesario distinguir dos tipos de conjuntos infinitos, los *numerables* y los *no numerables.* En esta sección se estudian ambos.

Para contar los elementos de un conjunto finito se pone en correspondencia el conjunto, elemento a elemento, con el conjunto de los números naturales $\{1, 2, \ldots, n\}$. La comparación de los «tamaños» de dos conjuntos mediante la correspondencia entre ellos elemento a elemento sustituye el recuento de los elementos cuando se trata de conjuntos infinitos. A este proceso se le puede dar una clara formulación matemática empleando el concepto de función:

DEFINICIÓN *Se dice que dos conjuntos A y B están en correspondencia uno a uno si existe una función f con las propiedades siguientes:*

a) *El dominio de f es A y el recorrido de f es B.*

b) *Si x e y son elementos distintos de A, entonces $f(x)$ y $f(y)$ son elementos distintos de B. Esto es, para todo par de elementos x e y de A,*

(13.21) $$x \neq y \quad \textit{implica} \quad f(x) \neq f(y).$$

Una función que satisfaga la propiedad (13.21) se llama *uno a uno* sobre A. Dos conjuntos A y B en correspondencia uno a uno se llaman también *equivalentes,* e indicamos esto poniendo $A \sim B$. Resulta claro que todo conjunto A es equivalente a sí mismo, puesto que $f(x) = x$ para cada x de A.
Un conjunto puede ser equivalente a un subconjunto de sí mismo. Por ejemplo el conjunto $P=\{1,2,3\ldots\}$, compuesto por todos los números enteros positivos, es equivalente a su subconjunto $Q=\{2,4,6\ldots\}$ compuesto por los pares positivos. En este caso, la función uno a uno que los hace equivalentes es $f(x)=2x$ para todo x de P.

Si $A \sim B$ es fácil de demostrar que $B \sim A$. Si f es uno a uno en A y si el recorrido de f es B, entonces para cada b en B existe excatamenet un a en A tal que $f(a) = b$. De ahí que podemos definir una función inversa g en B del modo siguiente: Si $b \in B$, $g(b) = a$, donde a es el único elemento de A tal que $f(a) = b$. Esta g es uno a uno en B y su recorrido es A; luego $B \sim A$. Esta propiedad de equivalencia se llama *simetría.*

(13.22) $$A \sim B \quad \text{implica} \quad B \sim A.$$

También es fácil demostrar que la equivalencia tiene la propiedad siguiente, llamada *transitividad:*

(13.23) $$A \sim B \quad \text{y} \quad B \sim C \quad \text{implica} \quad A \sim C.$$

En el ejercicio 2 de la sección (13.20) se propone una demostración de la propiedad transitiva.

Un conjunto S se denomina *finito* y se dice que contiene n elementos si

$$S \sim \{1, 2, \ldots, n\}.$$

El conjunto vacío también se considera finito. A los conjuntos que no son finitos se les llama *infinitos*. Un conjunto S se llama *infinito numerable* si es equivalente al conjunto de todos los números naturales, esto es, si

(13.24) $$S \sim \{1, 2, 3, \ldots\}.$$

En este caso existe una función f que establece una correspondencia uno a uno entre los enteros positivos y los elementos de S; luego el conjunto S puede expresarse según la notación en lista así:

$$S = \{f(1), f(2), f(3), \ldots\}.$$

A menudo utilizamos subíndices y representamos $f(k)$ con a_k (o con una notación parecida) y escribimos $S = \{a_1, a_2, a_3, \ldots, \}$. La idea importante es aquí que la correspondencia (13.24) nos permite usar los métodos naturales como «marcas» de los elementos de S.

Un conjunto se dice que es *numerable en sentido amplio* si es finito o infinito numerable. Un conjunto que no es numerable se llama *no numerable*. (Se darán ejemplos.) Muchas operaciones con conjuntos efectuadas sobre conjuntos numerables producen conjuntos numerables. Por ejemplo, tenemos las propiedades siguientes:

a) Todo subconjunto de un conjunto numerable es numerable.

b) La intersección de toda colección de conjuntos numerables es numerable.

c) La reunión de una colección numerable de conjuntos numerables es numerable.

d) El producto cartesiano de un número finito de conjuntos numerables es numerable.

Puesto que en este libro trataremos muy poco de los conjuntos infinitos numerables, no daremos con detalle las demostraciones de sus propiedades.(*) En cambio, ofrecemos varios ejemplos para poner de manifiesto cómo con ellas se pueden construir nuevos conjuntos numerables a partir de unos dados.

(*) En los ejercicios 3 al 8 de la sección 13.20 se esbozan las demostraciones.

EJEMPLO 1. El conjunto S de todos los números enteros (positivos, negativos y el cero) es numerable.

Demostración. Si $n \in S$, sea $f(n) = 2n$ si n es positivo, y $f(n) = 2|n| + 1$ si n es negativo o cero. El dominio de f es S y su recorrido es el conjunto de enteros positivos. Puesto que f es uno a uno en S, se deduce que S es numerable.

EJEMPLO 2. El conjunto R de todos los números racionales es numerable.

Demostración. Para cada entero $n \geqslant 1$ fijo, sea S_n el conjunto de números racionales de la forma x/n, donde x pertenece al conjunto S del ejemplo 1. Cada S_n es equivalente a S [tómese $f(t) = nt$ si $t \in S_n$] y por consiguiente cada S_n es numerable. Puesto que R es la reunión de todos los S_n, en virtud de la propiedad c) resulta R numerable.

Observación. Si $\mathscr{F} = \{A_1, A_2, A_3, \ldots\}$ es una colección numerable de conjuntos, la reunión de todos ellos se expresa con el símbolo

$$\bigcup_{k=1}^{\infty} A_k \qquad \text{o} \qquad A_1 \cup A_2 \cup A_3 \cup \cdots .$$

EJEMPLO 3. Sea $A = \{a_1, a_2, a_3, \ldots, \}$ un conjunto infinito numerable. Para cada entero $n \geqslant 1$, sea $\mathscr{F}_n$ la familia de subconjuntos de A con n elementos. Esto es:

$$\mathscr{F}_n = \{S \mid S \subseteq A \text{ y } S \text{ tiene } n \text{ elementos}\}.$$

Vamos a probar que cada $\mathscr{F}_n$ es numerable.

Demostración. Si S es un subconjunto de n elementos de A, podemos escribir:

$$S = \{a_{k_1}, a_{k_2}, \ldots, a_{k_n}\},$$

donde $k_1 < k_2 < \ldots < k_n$. Sea $f(S) = (a_{k_1}, a_{k_2}, \ldots, a_{k_n})$. Esto es, f es la función que asocia a S la n-pla ordenada $(a_{k_1}, a_{k_2}, \ldots, a_{k_n})$. El dominio de f es $\mathscr{F}_n$ y su recorrido, que designamos con T_n, es un subconjunto del producto cartesiano $C_n = A \times A \times \ldots \times A$ (n factores). Ya que A es numerable, lo mismo le ocurre a C_n [por la propiedad d)] y, por lo tanto, T_n también lo es [por la propiedad a)]. Pero $T_n \sim \mathscr{F}_n$ puesto que f es uno a uno. Esto demuestra que $\mathscr{F}_n$ es numerable.

EJEMPLO 4. La colección de todos los subconjuntos finitos de un conjunto numerable es numerable.

Demostración. El resultado es evidente si el conjunto dado es finito. Supon-

gamos, pues, que el conjunto dado (llamémosle A) es infinito numerable, y representemos con $\mathscr{F}$ la clase de todos los subconjuntos finitos de A:

$$\mathscr{F} = \{S \mid S \subseteq A \text{ y } S \text{ es finito}\}$$

En estas condiciones $\mathscr{F}$ es la reunión de todas las familias $\mathscr{F}_n$ del ejemplo 3; luego, en virtud de la propiedad c), $\mathscr{F}$ es numerable.

EJEMPLO 5. La colección de *todos* los subconjuntos de un conjunto infinito numerable es no numerable.

Demostración. Sea A el conjunto numerable dado y $\mathscr{A}$ la familia de todos los subconjuntos de A. Supondremos que $\mathscr{A}$ es numerable y llegaremos a una contradicción. Si $\mathscr{A}$ es numerable, entonces $\mathscr{A} \sim A$ y por tanto existe una función f uno a uno cuyo dominio es A y cuyo recorrido es $\mathscr{A}$. Así, para cada a de A, el valor $f(a)$ de la función es un subconjunto de A. Este subconjunto puede o no contener el elemento a. Designemos con B el conjunto de elementos a tales que $a \notin f(a)$. Así

$$B = \{a \mid a \in A \text{ pero } a \notin f(a)\}.$$

Siendo B un subconjunto de A, debe pertenecer a la familia $\mathscr{A}$. Esto significa que $B = f(b)$ para algún b de A. Existen ahora dos posibilidades: I) $b \in B$, o II) $b \notin B$. Si $b \in B$, según la definición de B tenemos $b \notin f(b)$, que es una contradicción, ya que $f(b) = B$. Por tanto I) es imposible. En el caso II), $b \notin B$, o sea, $b \notin f(b)$. Esto contradice la definición de B, con lo que II) también es imposible. Por consiguiente, el suponer que $\mathscr{A}$ es numerable nos lleva a una contradicción y debemos concluir que $\mathscr{A}$ es no numerable.

A continuación ofrecemos un ejemplo de conjunto no numerable más sencillo que el del ejemplo 5.

EJEMPLO 6. El conjunto de números reales x que satisfacen $0 < x < 1$ es no numerable.

Demostración. Supongamos otra vez que el conjunto es numerable y llegaremos a una contradicción. Si el conjunto es numerable podemos disponer sus elementos así: $\{x_1, x_2, x_3, \ldots\}$. Construiremos ahora un número real y que cumpla $0 < y < 1$ y que no estará en la lista. A tal fin escribimos cada elemento x_n en forma decimal:

$$x_n = 0.a_{n,1}\, a_{n,2}\, a_{n,3} \ldots,$$

donde cada $a_{n,i}$ es uno de los enteros del conjunto $\{0, 1, 2, \ldots, 9\}$. Sea y el número real cuyo desarrollo decimal es:

$$y = 0.y_1 y_2 y_3 \ldots,$$

donde,

$$y_n = \begin{cases} 1 & \text{si} \quad a_{n,n} \neq 1, \\ 2 & \text{si} \quad a_{n,n} = 1. \end{cases}$$

De este modo ningún elemento del conjunto $\{x_1, x_2, x_3, \ldots\}$ puede ser igual a y, puesto que y difiere de x_1 en la primera cifra decimal, de x_2 en la segunda, y en general, difiere de x_k en la k-ésima cifra decimal. Por tanto y satisface $0 < y < 1$, lo cual es una contradicción, lo que prueba que el conjunto de números reales del intervalo abierto (0,1) es no numerable.

13.20 Ejercicios

1. Sea $P = \{1, 2, 3, \ldots\}$ el conjunto de los números naturales. Para cada uno de los siguientes conjuntos, dar una función uno a uno f cuyo dominio es P y cuyo recorrido es el conjunto en cuestión:
a) $A = \{2, 4, 6, \ldots\}$, conjunto de los números pares positivos.
b) $B = \{3, 3^2, 3^3, \ldots\}$, conjunto de las potencias de 3.
c) $C = \{2, 3, 5, 7, 11, 13, \ldots\}$, conjunto de los números primos. [*Nota.* Una parte de la demostración consiste en demostrar que C es un conjunto infinito.]
d) $P \times P$, producto cartesiano de P por sí mismo.
e) El conjunto de enteros de la forma $2^m 3^n$, donde m y n son naturales.
2. Demostrar la propiedad transitiva de la equivalencia de conjuntos

$$\text{Si} \quad A \sim B \quad \text{y} \quad B \sim C, \quad \text{entonces} \quad A \sim C.$$

[*Indicación.* Si f hace A equivalente a B y g hace B equivalente a C, demostrar que la función compuesta $h = g \circ f$ hace A equivalente a C.]

Los ejercicios del 3 al 8 están dedicados a las demostraciones de las propiedades a), b), c), d) de los conjuntos numerables citadas en la sección 13.19.

3. Demostrar que todo subconjunto de un conjunto numerable es numerable. [*Indicación.* Supóngase S un conjunto infinito numerable, sea $S = \{x_1, x_2, x_3, \ldots\}$, y sea A un subconjunto infinito de S. Sea $k(1)$ el menor número natural m tal que $x_m \in A$. Admitamos que $k(1), k(2), \ldots, k(n-1)$ estén definidos, sea $k(n)$ el menor número natural $m > k(n-1)$ tal que $x_m \in A$. Sea $f(n) = x_{k(n)}$. Demostrar que f es una función uno a uno cuyo dominio es el conjunto de los números naturales y cuyo rango es A. Esto demuestra el teorema en el supuesto de que S sea infinito numerable. Construir otra demostración para S finito.]
4. Demostrar que la intersección de cualquier colección de conjuntos numerables es numerable. [*Indicación.* Utilizar el resultado del ejercicio 3.]
5. Sea $P = \{1, 2, 3, \ldots\}$ el conjunto de los números naturales.
a) Demostrar que el producto cartesiano $P \times P$ es numerable. [*Indicación.* Desígnese

con Q el conjunto de números naturales de la forma 2^m3^n, donde m y n son naturales. Entonces $Q \subset P$, de modo que Q es numerable (en virtud del ejercicio 3). Si $(m, n) \in P \times P$, tómese $f(m, n) = 2^m3^n$ y utilizar esta función para demostrar que $P \times P \sim Q$.]
b) Deducir de a) que el producto cartesiano de dos conjuntos numerables es numerable. Utilizar el método de inducción y extender el resultado a n conjuntos numerables.

6. Sea $\mathscr{B} = \{B_1, B_2, B_3, \ldots\}$ una colección numerable de conjuntos *disjuntos* ($B_i \cap B_j = \varnothing$ si $i \neq j$) tal que cada B_n es numerable. Demostrar que la reunión $\bigcup_{k=1}^{\infty} B_k$ es también numerable. [*Indicación.* Sea $B_n = \{b_{1,n}, b_{2,n}, b_{3,n}, \ldots\}$ y $S = \bigcup_{k=1}^{\infty} B_k$. Si $x \in S$, entonces $x = b_{m,n}$ para algún único par (m, n) y podemos definir $f(x) = (m, n)$. Utilizar esta f para demostrar que S es equivalente a un subconjunto de $P \times P$ y deducir (en virtud del ejercicio 5) que S es numerable.]

7. Sea $\mathscr{A} = \{A_1, A_2, A_3, \ldots\}$ una colección numerable de conjuntos, y definamos $\mathscr{B} = \{B_1, B_2, B_3, \ldots\}$ así: $B_1 = A_1$ y, para $n > 1$,
$$B_n = A_n - \bigcup_{k=1}^{n-1} A_k .$$
Esto es, B_n consta de los elementos de A_n que no pertenecen a los conjuntos precedentes $A_1, A_2, \ldots A_{n-1}$.
Demostrar que $\mathscr{B}$ es una colección de conjuntos disjuntos ($B_i \cap B_j = \varnothing$ si $i \neq j$) y que
$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k .$$
Esto nos permite expresar la reunión de cualquier colección numerable de conjuntos como reunión de una colección numerable de conjuntos *disjuntos*.

8. Si $\mathscr{F}$ es una colección numerable de conjuntos numerables, demostrar que la reunión de todos los conjuntos de $\mathscr{F}$ es numerable. [*Indicación.* Utilizar los ejercicios 6 y 7.]

9. Demostrar que los siguientes conjuntos son numerables:
a) El conjunto de todos los intervalos del eje real con extremos racionales.
b) El conjunto de todos los círculos del plano con radios racionales y centros de coordenadas racionales.
c) Cualquier conjunto de intervalos disjuntos de longitud positiva.

10. Demostrar que los siguientes conjuntos son no numerables:
a) El conjunto de los números irracionales del intervalo (0, 1).
b) El conjunto de todos los intervalos de longitud positiva.
c) El conjunto de todas las sucesiones cuyos términos son los enteros 0 y 1. (Recuérdese que una sucesión es una función cuyo dominio es el conjunto de los números naturales.)

13.21. Definición de probabilidad para espacios muestrales infinitos numerables

En esta sección se extiende la definición de probabilidad a espacios muestrales infinitos numerables. Sean S un conjunto infinito numerable y $\mathscr{B}$ un álgebra

de Boole de subconjuntos de S. Definamos una medida de probabilidad P en $\mathscr{B}$ como se hizo en el caso finito, excepto que exigiremos la aditividad numerable además de la aditividad finita. Esto es, para toda colección infinita numerable $\{A_1, A_2, \ldots\}$ de elementos de $\mathscr{B}$, exigimos que

$$(13.25)\quad P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \quad \text{si} \quad A_i \cap A_j = \varnothing \quad \text{siempre que } i \neq j.$$

Las funciones de conjunto de aditividad finita que satisfacen (13.25) se llaman de *aditividad numerable* (o completamente aditivas). Naturalmente, esta propiedad exige también suponer que la reunión numerable $A_1 \cup A_2 \cup A_3 \cup \cdots$ pertenece a $\mathscr{B}$ cuando cada A_k pertenezca a $\mathscr{B}$. No todas las álgebras de Boole tienen esta propiedad. Las que la tienen se llaman σ-álgebras de Boole. Un ejemplo es el álgebra booleana de todos los subconjuntos de S.

DEFINICIÓN DE PROBABILIDAD PARA ESPACIOS MUESTRALES INFINITOS NUMERALES. *Sea $\mathscr{B}$ una σ-álgebra de Boole cuyos elementos son subconjuntos de un conjunto infinito numerable S dado. Una función de conjunto P se llama medida de probabilidad en $\mathscr{B}$ si es no negativa, de aditividad numerable y satisface* $P(S) = 1$.

Cuando $\mathscr{B}$ es el álgebra de Boole de todos los subconjuntos de S, una función de probabilidad queda completamente determinada mediante sus valores para los subconjuntos de un solo elemento (tales valores se llaman probabilidades puntuales). Todo subconjunto A de S es finito o infinito numerable, y la probabilidad de A se calcula sumando las probabilidades puntuales para todos los elementos de A,

$$P(A) = \sum_{x \in A} P(x).$$

La suma del segundo miembro tiene un número finito de sumandos o es una serie absolutamente convergente.

El ejemplo que sigue muestra un ejemplo de una prueba con un espacio muestral infinito numerable.

EJEMPLO. Se lanza una moneda repetidamente hasta que el primer resultado vuelve a aparecer por segunda vez; entonces termina el juego.

Como espacio muestral tomamos la colección de todos los posibles juegos que pueden hacerse. Este conjunto puede expresarse como la reunión de dos conjuntos infinitos numerables A y B que son

$$A = \{TT, THT, THHT, THHHT, \ldots\} \text{ y } B = \{HH, HTH, HTTH, HTTTH, \ldots\}.$$

Designamos los elementos del conjunto A (en el orden en que se citan en la lista) con $a_0, a_1, a_2, \ldots,$ y los de B con $b_0, b_1, b_2, \ldots$. Podemos asignar arbitrariamente probabilidades puntuales no negativas $P(a_n)$ y $P(b_n)$ con tal que

$$\sum_{n=0}^{\infty} P(a_n) + \sum_{n=0}^{\infty} P(b_n) = 1 .$$

Por ejemplo, supongamos que la moneda tenga probabilidad p de mostrar cara (H) y probabilidad $q = 1 - p$ de mostrar cruz (T), siendo $0 < p < 1$. Entonces sería natural la asignación de probabilidades puntuales

$$P(a_n) = q^2 p^n \qquad \text{y} \qquad P(b_n) = p^2 q^n. \tag{13.26}$$

Tal asignación es aceptable porque tenemos

$$\sum_{n=0}^{\infty} P(a_n) + \sum_{n=0}^{\infty} P(b_n) = q^2 \sum_{n=0}^{\infty} p^n + p^2 \sum_{n=0}^{\infty} q^n = \frac{q^2}{1-p} + \frac{p^2}{1-q} = q + p = 1 .$$

Supongamos ahora que queremos saber la probabilidad de que el juego termine después de exactamente $n + 2$ lanzamientos. Éste es el suceso $\{a_n\} \cup \{b_n\}$, y su probabilidad es

$$P(a_n) + P(b_n) = q^2 p^n + p^2 q^n.$$

La probabilidad de que termine el juego por lo menos en $n + 2$ lanzamientos es

$$\sum_{k=0}^{n} P(a_k) + \sum_{k=0}^{n} P(b_k) = q^2 \frac{1 - p^{n+1}}{1 - p} + p^2 \frac{1 - q^{n+1}}{1 - q} = 1 - qp^{n+1} - pq^{n+1} .$$

13.22 Ejercicios

Los ejercicios de esta sección se refieren al ejemplo de la sección 13.21.

1. Utilicemos las probabilidades asignadas en la ecuación (13.26), designemos con $f_n(p)$ la probabilidad de que el juego finalice exactamente en $n + 2$ tiradas. Calcular los valores máximo y mínimo absolutos de $f_n(p)$ en el intervalo $0 \le p \le 1$ para cada uno de los valores $n = 0, 1, 2, 3$.
2. Demostrar que es aceptable cada una de las asignaciones de probabilidades puntuales siguientes.

a) $P(a_n) = P(b_n) = \dfrac{1}{2^{n+2}}$ para $n = 0, 1, 2, \ldots$.

b) $P(a_n) = P(b_n) = \dfrac{1}{(n+2)(n+3)}$ para $n = 0, 1, 2, \ldots$.

3. Calcular la probabilidad de que termine el juego antes de la quinta tirada, utilizando:
   a) las probabilidades puntuales de (13.26).
   b) las probabilidades puntuales del ejercicio 2 a).
   c) las probabilidades puntuales del ejercicio 2 b).
4. Calcular la probabilidad de que se necesite un número impar de tiradas para terminar el juego, utilizando:
   a) las probabilidades puntuales de (13.26).
   b) las probabilidades puntuales del ejercicio 2 a).
   c) las probabilidades puntuales del ejercicio 2 b).

13.23 Ejercicios variados sobre probabilidades

1. ¿Cuál es la probabilidad de conseguir un siete con dos d dos correctos?
2. Diez señores y sus esposas están sentados al azar en un banquete. Calcular la probabilidad de que un determinado señor esté sentado junto a su esposa si: a) están sentados en una mesa circular; b) están sentados en fila.
3. Un armario tiene dos cajones. El cajón número 1 contiene cuatro monedas de oro y dos de plata. El cajón número 2 contiene tres monedas de oro y tres de plata. Se abre un cajón al azar y de él se extrae una moneda al azar. Calcular la probabilidad de cada uno de los siguientes sucesos:
   a) Que sea abierto el cajón número 2 y se saque una moneda de plata.
   b) Que sea extraída una moneda de oro del cajón abierto.
4. Dos cartas se escogen sucesivamente en una baraja de 52 cartas, teniendo cada carta la misma probabilidad de ser elegida.
   a) ¿Cuál es la probabilidad de que por lo menos una sea una espada?
   Sin examinar las dos cartas escogidas se echan en una bolsa. Se extrae de la bolsa una de las cartas y se examina viéndose que no es una espada. (Cada carta tiene la misma probabilidad de ser extraída.)
   b) ¿Cuál es ahora la probabilidad de tener por lo menos una espada?
   La carta antes extraída de la bolsa se devuelve a la misma y las dos cartas se barajan. Nuevamente se extrae una carta y se examina. No hacemos comparación alguna para ver si la carta es la misma que antes se extrajo. Otra vez se devuelve la carta a la bolsa y se barajan. Esta operación se hace tres veces, incluyendo la de la parte b), y cada vez la carta examinada no es espada.
   c) ¿Cuál es el espacio muestral y la función de probabilidad de este experimento? ¿Cuál es la probabilidad de que una de las dos cartas originales sea espada?
5. Un hombre tiene diez monedas de un centavo, 9 corrientes y 1 con dos caras. Toma una moneda al azar y la lanza seis veces, y siempre obtiene cara. Calcular la probabilidad de que haya tomado la moneda de dos caras.
6. Demostrar que es imposible «lastrar» un par de dados de modo que todo resultado comprendido entre 2 y 12 tenga la misma probabilidad de presentarse.
7. Un estudiante de segundo año de Caltech tiene un despertador que sonará a las hora fijada con probabilidad 0,7. Si suena, le despertará a tiempo para llegar a su clase de matemáticas con probabilidad 0,8. Si no suena, la probabilidad de que llegue a su clase es 0,3. Calcular la probabilidad de que llegue a tiempo a clase.
8. Tres caballos A, B, y C están en una carrera. El suceso «A vence a B» se designará

simbólicamente escribiendo AB. El suceso «A vence a B el cual vence a C» se designará por ABC, etc. Se sabe que

$$P(AB) = \tfrac{2}{3}, \qquad P(AC) = \tfrac{2}{3}, \qquad P(BC) = \tfrac{1}{2},$$

y que

$$P(ABC) = P(ACB), \qquad P(BAC) = P(BCA), \qquad P(CAB) = P(CBA).$$

a) Calcular la probabilidad de que venza A.
b) Calcular la probabilidad de que venza B.
c) Calcular la probabilidad de que venza C.
d) ¿Son independientes los sucesos AB, AC y CB?

9. La fase final de un largo cálculo exige la adición de tres enteros a_1, a_2, a_3. Se supone que: a) los cálculos de a_1, a_2 y a_3 son estocásticamente independientes; b) en el cálculo de cada a, existe una misma probabilidad p de que sea correcto y la probabilidad de cometer un error igual a $+1$ es la misma que la de cometerlo igual a -1; c) no puede cometerse error mayor que $+1$ o menor que -1. Recordar la posibilidad de compensación de errores, y calcular la probabilidad de que la suma $a_1 + a_2 + a_3$ sea correcta

10. *El juego de «sale el non»*. Se supone que n personas lanzan cada una una moneda de manera simultánea e independiente, donde $n \geq 3$; y p es la probabilidad de obtener cara con cada una de las monedas. Calcular la probabilidad de que en una tirada dada se presente un «non»—, es decir, una persona cuya moneda muestra un resultado distinto del de todas las demás.

11. Se supone que n personas juegan a «sale el non» con monedas correctas (como se ha descrito en el ejercicio 10). Para un entero m, calcular la probabilidad de que son necesarias precisamente m tiradas para que haya un «non».

12. Supongamos que una experiencia compuesta $(S, \mathscr{B}, P)$está determinada por dos pruebas estocásticamente independientes $(S_1, \mathscr{B}_1, P_1)$ y $(S_2, \mathscr{B}_2, P_2)$ donde $S = S_1 \times S_2$ y

$$P(x, y) = P_1(x)P_2(y)$$

para cada (x, y) de S. El objeto de este ejercicio es establecer la fórmula

$$P(U \times V) = P_1(U)P_2(V) \tag{13.27}$$

para todo par de subconjuntos U de $\mathscr{B}_1$ y V de $\mathscr{B}_2$. Los espacios muestrales S_1 y S_2 se suponen finitos o infinitos numerables.

a) Comprobar que la igualdad (13.27) es verdadera cuando U y V son conjuntos de un solo elemento, y también cuando por lo menos uno de ellos es vacío.

Supongamos ahora que

$$U = \{u_1, u_2, \ldots, u_k\} \qquad \text{y} \qquad V = \{v_1, v_2, \ldots, v_m\}.$$

En tal caso $U \times V$ consta de km pares (u_i, v_j). Para cada $i = 1, 2, \ldots, k$ sea A_i el conjunto de m pares de $U \times V$ cuyo primer componente es u_i.

b) Demostrar que los A_i son conjuntos disjuntos cuya reunión es $U \times V$.

c) Demostrar que

$$P(A_i) = \sum_{j=1}^{m} P(u_i, v_j) = P_1(u_i)P_2(V).$$

d) Deducir a partir de b) y c) que

$$P(U \times V) = \sum_{i=1}^{k} P(A_i) = P_1(U)P_2(V).$$

14

CÁLCULO DE PROBABILIDADES

14.1 Definición de probabilidad para espacios muestrales no numerables

Un segmento rectilíneo se descompone en dos partes, con el punto de subdivisión elegido al azar. ¿Cuál es la probabilidad de que los dos fragmentos tengan la misma longitud? ¿Cuál es la probabilidad de que el mayor tenga exactamente longitud doble que la del pequeño? ¿Cuál es la probabilidad de que el mayor tenga longitud por lo menos doble que el menor? Éstos son ejemplos de problemas de probabilidad en los que el espacio muestral es no numerable ya que consta de todos los puntos del segmento. En esta sección se extiende la definición de probabilidad incluyendo los espacios muestrales no numerables.

Si siguiéramos el mismo proceso que para espacios muestrales numerables tendríamos que partir de un conjunto no numerable S cualquiera y una σ-álgebra de Boole $\mathscr{B}$ de subconjuntos de S y definir una medida de probabilidad que fuera una función de conjunto P no negativa completamente aditiva definida sobre $\mathscr{B}$ siendo $P(S) = 1$. Esto origina ciertas dificultades técnicas que no se presentan cuando S es numerable. Pretender la descripción de esas dificultades nos llevaría demasiado lejos. Las evitaremos imponiendo restricciones iniciales al conjunto S y al álgebra de Boole $\mathscr{B}$.

Primeramente, restringimos S a ser un subconjunto del eje real $\mathbf{R}$, o del n-espacio $\mathbf{R}^n$. Para el álgebra booleana $\mathscr{B}$ empleamos subconjuntos especiales de S que, en el lenguaje de la moderna teoría de la integración, se llaman subconjuntos de S *medibles*. No intentaremos describir el significado preciso de un conjunto medible; en cambio, mencionaremos algunas de las propiedades que posee la clase de los conjuntos medibles.

Consideremos primero subconjuntos de $\mathbf{R}$. Los subconjuntos medibles tienen las propiedades siguientes:

1) Si A es medible, asimismo lo es $\mathbf{R} - A$, complemento de A.

2) Si $\{A_1, A_2, A_3, \dots\}$ es una colección numerable de conjuntos medibles, la reunión $A_1 \cup A_2 \cup A_3 \cup \dots$ también es medible.

3) Todo intervalo (abierto, cerrado, semiabierto, finito o infinito) es medible.

Así pues, los conjuntos medibles de **R** forman una σ-álgebra booleana que contiene los intervalos. Existe una σ-álgebra booleana mínima (en el sentido de la inclusión) que tiene esta propiedad; sus elementos se llaman *conjuntos de Borel* en atención al matemático francés Émile Borel (1871-1956). Análogamente, en E_2 existe un σ-álgebra booleana mínima que contiene todos los productos cartesianos de pares de intervalos; sus elementos son conjuntos de Borel. Los conjuntos de Borel en E_n se definen en forma análoga.

De ahora en adelante, siempre que utilicemos un conjunto S de números reales como espacio muestral, o, más general, siempre que se use un conjunto S de E_n como espacio muestral, supondremos que este conjunto es un conjunto de Borel. Los subconjuntos de Borel de S forman asimismo una σ-álgebra de Boole; supondremos que nuestras medidas de probabilidad están definidas sobre esas álgebras de Boole. Dichas álgebras son lo bastante amplias para incluir todos los sucesos que se presentan en las aplicaciones corrientes de la teoría de probabilidades.

DEFINICIÓN DE PROBABILIDAD PARA ESPACIOS MUESTRALES NO NUMERABLES. *Sean S un subconjunto de $\mathbf{R}^n$, y $\mathscr{B}$ la σ-álgebra de Boole de subconjuntos de Borel de S. Una función de conjunto no negativa, completamente aditiva P definida en $\mathscr{B}$ con $P(S) = 1$ se llama medida de probabilidad. La terna $(S, \mathscr{B}, P)$ se llama espacio de probabilidad.*

14.2 Numerabilidad del conjunto de puntos con probabilidad positiva

Para espacios muestrales numerables la probabilidad de un suceso A se calcula a menudo sumando las probabilidades puntuales $P(x)$ para todo x de A. Este método es aplicable para espacios muestrales no numerables porque, como el siguiente teorema pone de manifiesto, la mayor parte de las probabilidades puntuales son nulas.

TEOREMA 14.1. *Si $(S,\mathscr{B},P)$ es un espacio de probabilidad y si T es el conjunto de todos los x de S para los que $P(x)>0$, entonces T es numerable.*

Demostración. Para cada $n = 1, 2, 3, \dots,$ designemos con T_n el siguiente subconjunto de S:

$$T_n = \left\{x \,\middle|\, \frac{1}{n+1} < P(x) \leq \frac{1}{n}\right\}.$$

Si $P(x) > 0$ entonces $x \in T_n$ para un cierto n. Recíprocamente, si $x \in T_n$ para un

cierto n entonces $x \in T$. Luego $T = T_1 \cup T_2 \cup \cdots$. Ahora bien T_n contiene a lo sumo n puntos, pues si contuviera $n + 1$ puntos o más la suma de sus probabilidades puntuales excedería a 1. Por lo tanto T es numerable, puesto que es una reunión de conjuntos finitos.

El teorema 14.1 nos dice que pueden asignarse probabilidades positivas a lo sumo a un subconjunto numerable de S. Los restantes puntos de S tendrán probabilidad cero. En particular, si todos los resultados de S son igualmente probables, entonces todo punto de S debe tener asignada probabilidad cero.

Observación: Puede darse una interpretación física del teorema 14.1 en función de la distribución de masa que ayuda a comprender su significado. Imaginemos que disponemos de una cantidad de masa que totaliza 1. (Esto corresponde a $P(S) = 1$.) Supongamos que nos es posible distribuir esa masa del modo que queramos a lo largo del eje real, bien con espesor uniforme o variable, o bien colocando porciones discretas de masa en ciertos puntos, o de ambas maneras. (Interpretamos una cantidad positiva de masa como una porción discreta.) Podemos colocar toda la masa en un punto. Podemos dividir la masa en porciones discretas iguales o desiguales entre dos puntos, diez puntos, un millón de puntos, o entre un conjunto no numerable de puntos. Por ejemplo, podemos poner $\frac{1}{2}$ en 1, $\frac{1}{4}$ en 2, $\frac{1}{8}$ en 3, etcétera, poner masa $(\frac{1}{2})^n$ en cada entero $n \geq 1$. O podemos esparcir la masa sin ninguna porción concentrada. O podemos esparcir parte de ella y distribuir el resto en porciones discretas. El teorema 14.1 nos dice que a lo más podemos asignar cantidades discretas de masa a un conjunto numerable de puntos.

Puesto que la mayor parte de (si no todas) las probabilidades puntuales para un espacio muestral no numerable será cero, no basta conocer las probabilidades puntuales para calcular las probabilidades de sucesos arbitrarios. Se necesita más información; la descripción es mejor utilizando dos conceptos nuevos, *las variables aleatorias* y las *funciones de distribución,* a las que nos vamos a referir seguidamente. Dichos conceptos permiten hacer uso del Cálculo integral en muchos problemas con espacios muestrales no numerables. La integración sustituye a la sumación en el Cálculo de probabilidades.

14.3 Variables aleatorias

En muchos experimentos nos interesan *números* asociados a los resultados del experimento. Por ejemplo, n monedas se lanzan simultáneamente y preguntamos por el número de caras. Un par de dados ruedan y nos interesa la suma de los puntos conseguidos. Se lanza una flecha hacia un blanco circular y queremos saber la distancia desde el punto en que cayó, al centro. Siempre que asociamos un número real a cada resultado de un experimento estamos tratando con una *función* cuyo dominio es el conjunto de resultados posibles y cuyo recorrido es el conjunto de los números reales en cuestión. Una tal función se llama *variable aleatoria.* A continuación damos una definición formal de variable aleatoria:

DEFINICIÓN DE VARIABLE ALEATORIA. *Sea S un espacio muestral. Una función real definida en S se llama variable aleatoria unidimensional. Si los valores de la función son pares ordenados de números reales (esto es, vectores en* E_2*), la función es una variable aleatoria bidimensional. Más general, una variable aleatoria n-dimensional es simplemente una función cuyo dominio es el espacio muestral dado S y cuyo recorrido es un colección de n-plas de números reales (vectores en* E_n*).*

Así pues, una variable aleatoria no es más que una función vectorial definida en un conjunto. La palabra «aleatoria» tan sólo se usa para recordar que el conjunto en cuestión es un espacio muestral (*).

A causa de la generalidad de la anterior definición es posible tener distintas variables aleatorias asociadas a un mismo experimento. En cada caso particular el experimentador debe decidir cuáles son las variables aleatorias que le interesan. Generalmente, procuramos trabajar con las variables aleatorias cuyas funciones reflejan, con la máxima simplicidad posible, las propiedades de los resultados del experimento que son realmente esenciales.

Notaciones. Ordinariamente se emplean para representar variables aleatorias unidimensionales letras mayúsculas tales como X, Y, Z. El resultado de un experimento se representa con la letra griega ω. Así que, $X(\omega)$ representa aquel número real que en la variable aleatoria X está asociado al resultado ω.

A continuación damos algunos ejemplos sencillos de variables aleatorias.

EJEMPLO 1. Un experimento consiste en echar un dado y leer el número de puntos conseguido. La variable aleatoria más «natural» a considerar es el número grabado por el fabricante en el dado, a saber:

$$X(\omega) = \omega \quad \text{para} \quad \omega = 1, 2, 3, 4, 5, 6.$$

Si estamos interesados en si el número de puntos es par o impar, debemos considerar otra variable aleatoria Y, definida como sigue:

$$Y(\omega) = 0 \quad \text{si} \quad \omega \text{ es par,}$$
$$Y(\omega) = 1 \quad \text{si} \quad \omega \text{ es impar.}$$

(*) La expresión «variable estocástica» se usa también como sinónimo de «variable aleatoria». La palabra «estocástica» procede de una raíz griega que significa «suerte» y según parece fue inventada por Jakob Bernoulli y se usa frecuentemente en la literatura de la teoría de probabilidades.

Los valores 0 y 1 no son esenciales —podría utilizarse cualquier par de números reales distintos. No obstante, 0 y 1 sugieren «par» o «impar», respectivamente, porque representan el resto que se obtiene al dividir el resultado ω por 2.

EJEMPLO 2. Se lanza una flecha hacia un blanco circular. El conjunto de todos los resultados posibles es el de todos los puntos ω del blanco. Si imaginamos un sistema de coordenadas colocado en el blanco con el origen en su centro, podemos asignar diversas variables aleatorias a este experimento. Una de ellas es la variable bidimensional que asigna al punto ω sus coordenadas rectangulares (x, y). Otra es el par de coordenadas polares (r, θ) del punto ω. Como ejemplos de variables unidimensionales son aquellas que asignan a cada ω una de sus coordenadas, tales como la x o la r. En un experimento de este tipo deseamos a menudo conocer la probabilidad de que la flecha toque en una determinada región del blanco, por ejemplo, el primer cuadrante. Este suceso se puede describir de la manera más simple mediante la variable aleatoria que asigna a cada punto ω su coordenada polar θ, así que $X(\omega) = \theta$; el suceso «la flecha da en el primer cuadrante» es el conjunto de valores de ω tales que $0 \leq X(\omega) \leq \frac{1}{2}\pi$.

Notaciones. Evitamos la notación engorrosa empleando notaciones brëves para describir ciertos tipos de sucesos y sus probabilidades. Por ejemplo, si t es un número real, el conjunto de todos los ω del espacio muestral tales que $X(\omega) = t$ se representa brevemente poniendo

$$X = t.$$

La probabilidad de ese suceso se nota $P(X = t)$ en lugar de poner $P(\{\omega | X(\omega) = t\})$. Los símbolos tales como $P(X = a \text{ o } X = b)$ y $P(a < X \leq b)$ se definen de modo parecido. Así pues, el suceso «$X = a$ o $X = b$» es la reunión de los dos sucesos «$X = a$» y «$X = b$»; el símbolo $P(X = a \text{ o } X = b)$ representa la probabilidad de esa reunión. El suceso «$a < X \leq b$» es el conjunto de todos los puntos ω tales que $X(\omega)$ pertenece al intervalo semiabierto $(a, b]$, y el símbolo $P(a < X \leq b)$ representa la probabilidad de ese suceso.

14.4 Ejercicios

1. Sea X una variable aleatoria unidimensional.
   a) Si $a < b$, demostrar que los dos sucesos $a < X \leq b$ y $X \leq a$ son disjuntos.
   b) Determinar la reunión de los dos sucesos de la parte a).
   c) Demostrar que $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$.
2. Sea (X, Y) una variable aleatoria bidimensional definida sobre un espacio muestral S. Esto significa que (X, Y) es una función que asigna a cada ω de S un par de números reales $(X(\omega), Y(\omega))$. Naturalmente, cada X y cada Y es una variable aleatoria unidimensional definida en S. La notación

$$X \leq a, Y \leq b$$

representa el conjunto de todos los elementos ω de S tales que $X(\omega) \leq a$ y $Y(\omega) \leq b$.
a) Si $a < b$ y $c < d$, expresar, en función de los elementos de S, el significado de la siguiente notación: $a < X \leq b$, $c < Y \leq d$.
b) Demostrar que los dos sucesos «$X \leq a, Y \leq c$» y «$X \leq a, c < Y \leq d$» son disjuntos. Interpretar geométricamente esos sucesos.
c) Determinar la reunión de los dos sucesos del apartado b).
d) Generalizar el ejercicio 1 c) al caso bidimensional.

3. Se echan dos dados correctos, siendo cada resultado un par ordenado (a, b), siendo a y b enteros del 1 al 6. Sea X la variable aleatoria que asigna el valor $a + b$ al resultado (a, b).
a) Describir, en la notación en lista, los sucesos «$X = 7$», «$X = 11$», «$X = 7$ ó $X = 11$».
b) Calcular las probabilidades de los sucesos de la parte a).
4. Consideremos un experimento en el que se lanzan cuatro monedas simultáneamente (o una moneda se lanza cuatro veces). Para cada moneda definimos una variable aleatoria que asigna el valor 1 a la cara y el 0 a la cruz, y representemos esas variables aleatorias por X_1, X_2, X_3, X_4, asignando las probabilidades $P(X_i = 1) = P(X_i = 0) = \frac{1}{2}$ para cada X_i Considerar una nueva variable aleatoria Y que asigne a cada resultado el número total de caras de las cuatro monedas. Expresar Y en función de X_1, X_2, X_3, X_4 y calcular las probabilidades $P(Y = 0)$, $P(Y = 1)$, y $P(Y \leq 1)$.
5. Una pequeña compañía ferroviaria tiene medios para el transporte de 100 pasajeros por día entre dos ciudades, por un coste (para la compañía) de 7 $ por pasajero. Si algún día compran billetes más de 100 pasajeros la compañía está obligada a proporcionar transporte en autobús a los pasajeros que exceden por un coste de 10 $ por persona. Sea X la variable aleatoria que representa el número de pasajeros que compran billetes en un día determinado. Los valores posibles de X son los enteros 0, 1, 2, 3, ... hasta una cierto máximo desconocido. Si con Y representamos la variable aleatoria que expresa el coste total diario (en dólares) para la compañía por los pasajeros despachados, expresar Y en función de X.
6. Una factoría de producción en cadena consta de dos estaciones de trabajo A y B. En la estación A, se montan X unidades por hora; inmediatamente son transportadas a la estación B, donde son inspeccionadas a razón de Y unidades por hora, siendo $Y < X$. Los posibles valores de X e Y son los enteros 8, 9 y 10. Si Z es la variable aleatoria que representa el número de unidades que salen de la cadena en la primera hora de producción:
a) Expresar Z en función de X e Y, suponiendo cada X y cada Y constantes durante esta hora.
b) Describir, de manera parecida, la variable aleatoria U que cuenta el número de unidades producidas en las dos primeras horas. Cada una de las X e Y es constante durante cada hora, pero los valores constantes durante la segunda hora no deben necesariamente coincidir con los de la primera.

14.5 Funciones de distribución

Volvemos de nuevo al problema del cálculo de probabilidades de un suceso asociado a una cierta variable aleatoria. Sea X una variable aleatoria unidimensional definida en un espacio muestral S, siendo S un conjunto de Borel en E_n para $n \geq 1$. Sea P una medida de probabilidad definida en los subconjuntos de Borel de S. Para cada ω de S, $X(\omega)$ es un número real, y cuando ω recorre los elementos de S los números $X(\omega)$ recorren un conjunto de números reales (el recorrido de X).

Este conjunto puede ser finito, infinito numerable, o no numerable. Para cada número real t consideramos el siguiente subconjunto de S:

$$A(t) = \{\omega \mid X(\omega) \leq t\}.$$

Si t es menor que todos los números del recorrido de X, el conjunto $A(t)$ será vacío; de lo contrario, $A(t)$ será un subconjunto de S no vacío. Suponemos que para cada t el conjunto $A(t)$ es un *suceso,* esto es, un conjunto de Borel. De acuerdo con el convenio establecido al final de la sección 14.3, representamos este suceso con el símbolo $X \leq t$.

Suponemos conocida la probabilidad $P(X \leq t)$ para todo t real. Este conocimiento nos permitirá calcular las probabilidades de otros muchos sucesos de interés. Para lograrlo se usan las probabilidades $P(X \leq t)$ como base para la construcción de una nueva función F, llamada *función de distribución* de X que se define de la manera siguiente:

DEFINICIÓN DE FUNCIÓN DE DISTRIBUCIÓN. *Sea X una variable aleatoria unidimensional. La función F definida para todo t real mediante la ecuación*

$$F(t) = P(X \leq t)$$

se llama la función de distribución de la variable aleatoria X.

Observación. Algunas veces, la notación F_X se utiliza para resaltar el hecho de que la función de distribución está asociada a una cierta variable X. Entonces el valor de la función en t se escribe $F_X(t)$.

Es importante darse cuenta de que la función de distribución F está definida sobre el eje real completo, aun cuando el recorrido de X puede ser una porción acotada del eje real. En efecto, si todos los números $X(\omega)$ están en un cierto intervalo finito $[a, b]$, entonces para $t < a$ la probabilidad $P(X \leq t)$ es cero (ya que para $t < a$ el conjunto $X \leq t$ es vacío) y para $t \geq b$ la probabilidad $P(X \leq t)$ es 1 (debido a que en este caso el conjunto $X \leq t$ es el espacio muestral completo). Esto significa que para variables aleatorias X *acotadas* cuyo recorrido está dentro de un intervalo $[a, b]$ tenemos $F(t) = 0$ para todo $t < a$ y $F(t) = 1$ para todo $t \geq b$.

Procederemos ahora a deducir un cierto número de propiedades comunes a todas las funciones de distribución.

TEOREMA 14.2. *Si F es una función de distribución de una variable aleatoria unidimensional X, entonces*

a) $0 \leq F(t) \leq 1$ *para todo t.*

b) $P(a < X \leq b) = F(b) - F(a)$ *si* $a < b$.

c) $F(a) \leq F(b)$ *si* $a < b$.

Demostración. La parte a) se deduce inmediatamente de la definición de F porque la probabilidad está siempre comprendida entre 0 y 1.

Para demostrar b) observemos que los sucesos «$a < X \leq b$» y «$X \leq a$» son disjuntos. Su reunión es el suceso «$X \leq b$». Aplicando la aditividad obtenemos

$$P(a < X \leq b) + P(X \leq a) = P(X \leq b),$$

que también se puede poner en la forma

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a).$$

La parte c) se deduce de b) puesto que $P(a < X \leq b) \geq 0$.

Observación: Usando la analogía de la masa, diríamos que $F(t)$ representa la cantidad total de masa situada entre $-\infty$ y t (incluyendo el mismo punto t). La cantidad de masa localizada en el intervalo semiabierto $(a, b]$ es $F(b) - F(a)$.

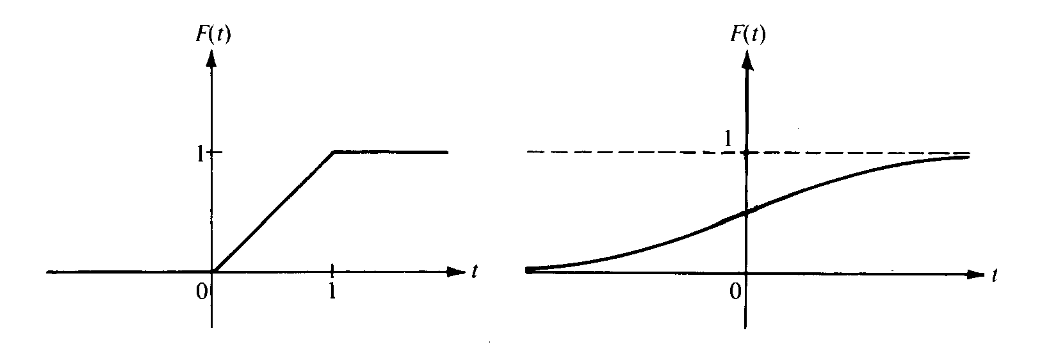


FIGURA 14.1 *Función de distribución de una variable aleatoria acotada.*

FIGURA 14.2 *Función de distribución de una variable aleatoria no acotada.*

La figura 14.1 nos muestra una función de distribución de una variable aleatoria acotada X cuyos valores $X(\omega)$ están en el intervalo $[0, 1]$. Este ejemplo es conocido como *distribución uniforme.* En él tenemos

$$F(t) = 0 \text{ para } t < 0, \qquad F(t) = t \text{ para } 0 < t \leq 1, \qquad F(t) = 1 \text{ para } t \geq 1.$$

En la figura 14.2 se ve un ejemplo de una función de distribución correspondiente a una variable aleatoria no acotada. Este ejemplo se conoce con el nombre de *distribución de Cauchy* y sus valores vienen dados por la fórmula

$$F(t) = \frac{1}{2} + \frac{1}{\pi} \arctan t.$$

Más adelante veremos casos prácticos que conducen a distribuciones uniformes y de Cauchy.

Observación: Con el símil de la masa, diríamos que en la figura 14.1 no existe masa situada a la izquierda del origen o a la derecha del punto 1. Toda la masa ha sido distribuida en el intervalo $[0, 1]$. La gráfica de F es una recta sobre ese intervalo porque la masa está extendida con densidad uniforme. En la figura 14.2 la masa está distribuida a lo largo de todo el eje. La gráfica no es lineal porque la masa está extendida con densidad variable.

El teorema 14.2 b) indica cómo calcular (en función de F) la probabilidad de que X esté en un intervalo semi-abierto de la forma $(a, b]$. El teorema que sigue se refiere a otros tipos de intervalos.

TEOREMA 14.3. *Si F es una función de distribución de una variable aleatoria uni-dimensional X, entonces si $a<b$ tenemos:*

a) $P(a \leq X \leq b) = F(b) - F(a) + P(X = a)$.
b) $P(a < X < b) = F(b) - F(a) - P(X = b)$.
c) $P(a \leq X < b) = F(b) - F(a) + P(X = a) - P(X = b)$.

Demostración. Para demostrar a) observemos que los sucesos «$a < X \leq b$» y «$X = a$» son disjuntos y su reunión es «$a \leq X \leq b$». Teniendo en cuenta la aditividad y el teorema 14.2 b) obtenemos a). Las partes b) y c) se demuestran en forma parecida.

Observemos que los cuatro sucesos

$$a < X \leq b, \qquad a \leq X \leq b, \qquad a < X < b, \qquad \text{y} \qquad a \leq X < b$$

tienen igual probabilidad si y sólo si $P(X = a) = 0$ y $P(X = b) = 0$.

Los ejemplos representados en las figuras 14.1 y 14.2 hacen ver otras dos propiedades que tienen todas las funciones de distribución, y que se describen en el teorema siguiente.

TEOREMA 14.4. *Si F es la función de distribución de una variable aleatoria uni-dimensional X, entonces*

$$(14.1) \qquad \lim_{t \to -\infty} F(t) = 0 \qquad y \qquad \lim_{t \to +\infty} F(t) = 1 .$$

Demostración. La existencia de los dos límites de (14.1) y el hecho de que los dos límites estén entre 0 y 1 se deduce al momento, puesto que F es una función monótona cuyos valores están entre 0 y 1.

Designemos los límites de (14.1) con L_1 y L_2 respectivamente. Para demos-

trar que $L_1 = 0$ y que $L_2 = 1$ emplearemos la propiedad de aditividad numerable de la probabilidad. A tal fin expresemos la totalidad del espacio S como una reunión numerable de sucesos disjuntos:

$$S = \bigcup_{n=1}^{\infty}(-n < X \leq -n + 1) \cup \bigcup_{n=0}^{\infty}(n < X \leq n + 1).$$

Entonces, por la aditividad, obtenemos

$$\begin{aligned} P(S) &= \sum_{n=1}^{\infty} P(-n < X \leq -n + 1) + \sum_{n=0}^{\infty} P(n < X \leq n + 1) \\ &= \lim_{M\to\infty} \sum_{n=1}^{M} [F(-n + 1) - F(-n)] + \lim_{N\to\infty} \sum_{n=0}^{N} [F(n + 1) - F(n)]. \end{aligned}$$

Las sumas del segundo miembro se simplifican unas con otras, dándonos

$$\begin{aligned} P(S) &= \lim_{M\to\infty} [F(0) - F(-M)] + \lim_{N\to\infty} [F(N + 1) - F(0)] \\ &= F(0) - L_1 + L_2 - F(0) = L_2 - L_1 . \end{aligned}$$

Puesto que $P(S) = 1$, esto demuestra que $L_2 - L_1 = 1$ o sea $L_2 = 1 + L_1$. Por otra parte, tenemos también que $L_2 \leq 1$ y $L_1 \geq 0$. Esto implica que $L_2 = 0$ y $L_1 = 1$, como se afirmó.

14.6 Discontinuidades de las funciones de distribución

En la figura 14.3 se muestra un ejemplo de una posible función de distribución con discontinuidades. Utilizando la analogía de masas diríamos que F tiene una discontinuidad de salto en cada punto que soporta una cantidad positiva de masa. El teorema que sigue muestra que el salto es igual a la cantidad de masa concentrada en el punto correspondiente.

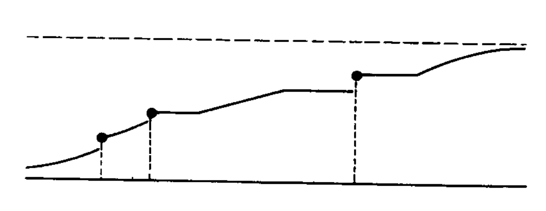


FIGURA 14.3 *Una posible función de distribución.*

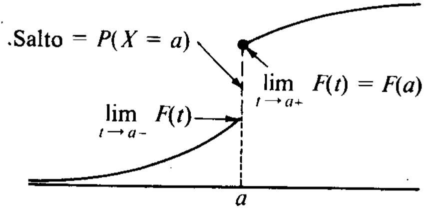


FIGURA 14.4 *Discontinuidad de salto de una función de distribución.*

TEOREMA 14.5. *Si F es una función de distribución de una variable aleatoria uni-dimensional X, entonces para cada valor real t tenemos*

(14.2) $$\lim_{t\to a+} F(t) = F(a)$$

y

(14.3) $$\lim_{t\to a-} F(t) = F(a) - P(X = a).$$

Observación: La relación entre límites (14.2) nos dice que F es *continua a la derecha* de cada punto a, debido a que $F(t) \to F(a)$ cuando $t \to a$ por la derecha. Por otra parte, la ecuación (14.3) nos dice que cuando $t \to a$ por la izquierda, $F(t)$ tiende hacia $F(a)$ si y sólo si la probabilidad $P(X = a)$ es cero. Cuando $P(X = a)$ no es cero, la gráfica de F tiene una discontinuidad de salto en a del tipo de la que se ve en la figura 14.4.

Demostración. La existencia de los límites resulta inmediatamente de la monotonía y acotación de F. Vamos a demostrar que los límites tienen los valores indicados. A tal i n utilizamos la parte b) del teorema 14.2. Si $t > a$ escribimos:

(14.4) $$F(t) = F(a) + P(a < X \leq t);$$

si $t < a$, ponemos

(14.5) $$F(t) = F(a) - P(t < X \leq a).$$

Haciendo $t \to a+$ en (14.4) encontramos

$$\lim_{t\to a+} F(t) = F(a) + \lim_{t\to a+} P(a < X \leq t),$$

mientras que si $t \to a-$ en (14.5) obtenemos

$$\lim_{t\to a-} F(t) = F(a) - \lim_{t\to a-} P(t < X \leq a).$$

Por lo tanto para demostrar (14.2) y (14.3) debemos establecer dos igualdades:

(14.6) $$\lim_{t\to a+} P(a < X \leq t) = 0$$

y

(14.7) $$\lim_{t\to a-} P(t < X \leq a) = P(X = a).$$

Intuitivamente pueden establecerse así: Cuando $t \to a+$, el intervalo semiabierto $(a, t]$ tiende a confundirse con el conjunto vacío. Esto es, la intersección de todos los intervalos semiabiertos $(a, t]$, para $t > a$, es vacía. Por otra parte, cuando $t \to a-$ el intervalo semiabierto $(t, a]$ tiende a confundirse con el punto a. (La intersección de todos los intervalos $(t, a]$ para $t < a$ es el conjunto $\{a\}$.) Por consiguiente, si la probabilidad se comporta de manera continua, las igualdades (14.6) y (14.7) deben ser ciertas. Para seguir este razonamiento en forma rigurosa vamos a hacerlo del siguiente modo:

Para cada entero $n \geq 1$, sea

$$p_n = P\left(a < X \leq a + \frac{1}{n}\right). \tag{14.8}$$

Para demostrar (14.6) basta con demostrar que $p_n \to 0$ cuando $n \to \infty$. Representando con S_n el suceso

$$a + \frac{1}{n+1} < X \leq a + \frac{1}{n}.$$

Los conjuntos S_n son disjuntos y su reunión $S_1 \cup S_2 \cup S_3 \cup \cdots\cdots$ es el suceso $a < X \leq a + 1$. Por la aditividad numerable tenemos:

$$\sum_{n=1}^{\infty} P(S_n) = P(a < X \leq a + 1) = p_1. \tag{14.9}$$

Por otra parte, la igualdad (14.8) implica que

$$p_n - p_{n+1} = P(S_n),$$

así que de (14.9) obtenemos la relación

$$\sum_{n=1}^{\infty} (p_n - p_{n+1}) = p_1. \tag{14.10}$$

La convergencia de la serie es una consecuencia de (14.9). Pero la serie del primer miembro de (14.10) tiene por suma

$$p_1 - \lim_{n \to \infty} p_n.$$

Por lo tanto (14.10) implica que $\lim_{n\to\infty} p_n = 0$, y esto prueba (14.6).

Una ligera modificación de este razonamiento nos permite demostrar también (14.7). Puesto que

$$P(t < X \leq a) = P(t < X < a) + P(X = a)$$

tan sólo necesitamos demostrar que

$$\lim_{t \to a-} P(t < X < a) = 0 .$$

A tal fin introducimos los números

$$q_n = P\left(a - \frac{1}{n} < X < a\right)$$

y demostramos que $q_n \to 0$ cuando $n \to \infty$. En este caso consideramos los sucesos T_n dados por

$$a - \frac{1}{n} < X \leq a - \frac{1}{n+1}$$

para $n = 1, 2, 3, \ldots$. Esos son disjuntos y su reunión es el suceso $a - 1 < X < a$, así que tenemos

$$\sum_{n=1}^{\infty} P(T_n) = P(a - 1 < X < a) = q_1 .$$

Observemos ahora que $q_n - q_{n+1} = P(T_n)$, y podemos completar la demostración como antes.

El tipo más general de distribución es cualquier función real F con las propiedades siguientes:

a) F es monótona creciente en el eje real.
b) F es continua a la derecha de cada punto.
c) $\lim_{t \to -\infty} F(t) = 0$ y $\lim_{t \to +\infty} F(t) = 1$.

En efecto, puede demostrarse que para cada una de esas funciones F existe una correspondiente función de conjunto P, definida sobre los conjuntos de Borel del eje real, tal que P es una medida de probabilidad que asigna la probabilidad $F(b) - F(a)$ a cada intervalo semiabierto $(a, b]$. Puede verse una demostración de esto en H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N. J., 1946.

Existen dos tipos especiales de distribuciones, llamadas *discretas* y *continuas*, que en la práctica tienen particular importancia. En el caso discreto toda la masa está concentrada en un número de puntos finito o infinito numerable, mientras que en el continuo la masa está esparcida, con espesor uniforme o variado, a lo largo de todo el eje real. En las próximas secciones se tratan con cierto detalle esos dos tipos de distribuciones.

14.7 Distribuciones discretas. Funciones de masa de probabilidad

Sea X una variable aleatoria uni-dimensional y consideremos una nueva función p, llamada *función de masa de probabilidad* de X. Sus valores $p(t)$ están definidos para todo número real t mediante la ecuación

$$p(t) = P(X = t).$$

Esto es, $p(t)$ es la probabilidad de que X tome el valor t. Cuando deseemos poner de manifiesto que p está asociado a X escribimos p_X en lugar de p y $p_X(t)$ en lugar de $p(t)$.

El conjunto de números reales t para los cuales $p(t) > 0$ es finito o numerable. Lo expresamos por T; esto es, ponemos

$$T = \{t \mid p(t) > 0\}.$$

La variable aleatoria X se llama *discreta* si

$$\sum_{t \in T} p(t) = 1.$$

Dicho de otro modo, X es discreta si una unidad de masa de probabilidad está distribuida sobre el eje real concentrándose una masa positiva $p(t)$ en cada punto t de un cierto conjunto T finito o infinito numerable y en los restantes puntos no hay masa. Los puntos de T se llaman *puntos de masa* de X.

Para variables aleatorias discretas el conocimiento de la función de masa de probabilidad nos permite calcular la probabilidad de sucesos arbitrarios. Tenemos, efectivamente, el siguiente teorema.

TEOREMA 14.6. *Si A es un subconjunto de Borel de la recta real* **R**, *y si designamos con $P(X \in A)$ la probabilidad del conjunto de ω tal que $X(\omega) \in A$, entonces*

$$P(X \in A) = \sum_{x \in A \cap T} p(x), \tag{14.11}$$

donde T es el conjunto de puntos de masa de X.

Demostración. Puesto que $A \cap T \subseteq A$ y $T - A \subseteq \mathbf{R} - A$,

$$\sum_{x \in A \cap T} p(x) \leq P(X \in A) \quad \text{y} \quad \sum_{x \in T - A} p(x) \leq P(X \notin A). \tag{14.12}$$

Pero $A \cap T$ y $T - A$ son conjuntos disjuntos cuya reunión es T, así que la segunda desigualdad de (14.12) es equivalente a

$$1 - \sum_{x \in A \cap T} p(x) \leq 1 - P(X \in A) \qquad \text{o} \qquad \sum_{x \in A \cap T} p(x) \geq P(X \in A).$$

Combinando esto con la primera desigualdad (14.12) obtenemos (14.11).

Observación: Puesto que $p(x) = 0$ cuando $x \notin T$, la suma del segundo miembro de (14.11) puede ponerse en la forma $\sum_{x \in A} p(x)$ sin peligro de confusión.

Cuando A es el intervalo $(-\infty, t]$, la suma de (14.11) da el valor de la función de distribución $F(t)$. Así pues, tenemos

$$F(t) = P(X \leq t) = \sum_{x \leq t} p(x).$$

Si una variable aleatoria es discreta, la función de distribución correspondiente F también es discreta.

Los ejemplos que siguen de distribuciones discretas se presentan frecuentemente en la práctica.

EJEMPLO 1. *Distribución binomial.* Sea p un número real que satisface la acotación $0 \leq p \leq 1$ y sea $q = 1 - p$. Supongamos una variable aleatoria que toma los valores $0, 1, 2, 3, \ldots, n$, siendo n un entero positivo fijo, y admitamos que la probabilidad $P(X = k)$ viene dada por la fórmula

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \qquad \text{para} \quad k = 0, 1, 2, \ldots, n.$$

a) La función masa. b) La función de distribución.

FIGURA 14.5 *Las funciones de masa de probabilidad y de distribución de una variable aleatoria discreta con una infinidad numerable de puntos de masa* $n = 5$ *y* $p = \frac{1}{3}$.

Esta asignación de probabilidades es legítima porque la suma de todas ellas es:

$$\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1 .$$

La correspondiente función de distribución F_X se denomina *distribución binomial* de parámetros n y p. Sus valores pueden calcularse mediante la fórmula de sumación

$$F_X(t) = \sum_{0 \le k \le t} \binom{n}{k} p^k q^{n-k} .$$

La distribución binomial se presenta de manera natural en el caso de una sucesión de pruebas de Bernoulli si p es la probabilidad de un suceso y q la del suceso contrario. Efectivamente, cuando la variable aleatoria X cuenta el número de veces que se presenta el suceso en n pruebas, $P(X = k)$ es precisamente $\binom{n}{k} p^k q^{n-k}$ en virtud de la fórmula de Bernoulli. (Véase el teorema 13.3 de la sección 13.16.) La figura 14.5 muestra las gráficas de la función de masa de probabilidad y la función de distribución correspondiente para una distribución binomial con parámetros $n = 5$ y $p = 1/3$.

EJEMPLO 2. *Distribución de Poisson.* Sean λ un número real positivo y X una variable aleatoria que toma los valores 0, 1, 2, 3, ... Si la probabilidad $P(X = k)$ está dada por la fórmula

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad \text{para} \quad k = 0, 1, 2, \ldots ,$$

la función de distribución correspondiente F_X se llama *distribución de Poisson* de parámetro λ. Se denomina así en honor del matemático francés S. D. Poisson (1781-1840). Esta asignación de probabilidades es válida porque

$$\sum_{k=0}^{\infty} P(X = k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1 .$$

Los valores de la función de distribución se calculan con las sumas parciales

$$F_X(t) = e^{-\lambda} \sum_{0 \le k \le t} \frac{\lambda^k}{k!} .$$

La distribución de Poisson es aplicable a diversos problemas relativos a sucesos aleatorios que se presentan en el transcurso del tiempo, tales como acciden-

tes de tráfico, conexiones equivocadas en una central telefónica, e intercambios de cromosomas en las células provocados por rayos X. En los libros de Feller y Parzen, que al final del capítulo se citan, se discuten algunas aplicaciones particulares.

14.8 Ejercicios

1. Se lanza un dado correcto. Como variable aleatoria X tomamos la función que cuenta el número de puntos conseguidos con el dado. Dibujar la gráfica de la correspondiente función de distribución F_X.
2. Se lanzan dos dados. Sea X la variable aleatoria que cuenta el total de puntos conseguidos. Construir una tabla que dé los valores no nulos de la función de masa de probabilidad p_X y dibujar la gráfica de la función de distribución F_X.
3. La función de distribución F de una variable aleatoria X viene definida como sigue

$$F(t) = \begin{cases} 0 & \text{si } t < -2, \\ \frac{1}{2} & \text{si } -2 \leq t < 0, \\ \frac{3}{4} & \text{si } 0 \leq t < 2, \\ 1 & \text{si } t \geq 2. \end{cases}$$

a) Dibujar la gráfica de F.
b) Describir la función de masa de probabilidad p y representarla.
c) Calcular las siguientes probabilidades: $P(X = 1)$, $P(X \leq 1)$, $P(X < 1)$, $P(X = 2)$ $P(X \leq 2)$, $P(0 < X < 2)$, $P(0 < X \leq 2)$, $P(1 \leq X \leq 2)$.

4. Considérese una variable aleatoria X cuyos posibles valores son todos los números racionales de la forma $\frac{n}{n+1}$ y $\frac{n+1}{n}$ siendo $n = 1, 2, 3, \ldots$. Si

$$P\left(X = \frac{n}{n+1}\right) = P\left(X = \frac{n+1}{n}\right) = \frac{1}{2^{n+1}},$$

comprobar que esta asignación de probabilidades es correcta y diseñar la forma de la gráfica de la función de distribución F_X.

5. La función de masa de probabilidad p de una variable aleatoria X es nula salvo en los puntos $t = 0, 1, 2$. En ellos toma los valores

$$p(0) = 3c^3, \qquad p(1) = 4c - 10c^2, \qquad p(2) = 5c - 1,$$

para un cierto $c > 0$.
a) Determinar el valor c.
b) Calcular las probabilidades: $P(X < 1)$, $P(X < 2)$, $P(1 < X \leq 2)$, $P(0 < X < 3)$.
c) Describir y representar la función de distribución F.
d) Hallar el mayor valor de t tal que $F(t) < \frac{1}{2}$.
e) Hallar el menor valor de t tal que $F(t) > \frac{1}{3}$.

6. Una variable aleatoria X tiene distribución binomial con parámetros $n = 4$ y $p = \frac{1}{3}$.
   a) Describir y representar la función de masa de probabilidad p.
   b) Describir y representar la función de distribución F.
   c) Calcular las probabilidades $P(1 < X \leq 2)$ y $P(1 \leq X \leq 2)$.
7. Supongamos que una chincheta se echa sobre la mesa, y cae o con la punta hacia arriba o en una posición estable con la punta en la mesa. Supongamos que existe una probabilidad p de que quede con la punta hacia arriba.
   a) Supongamos que se echan en la mesa dos chinchetas idénticas. Admitiendo la independencia estocástica, demostrar que la probabilidad de que ambas queden con la punta hacia arriba es p^2.
   b) Continuando la parte a), sea X la variable aleatoria que cuenta el número de chinchetas que quedan con la punta hacia arriba (los valores posibles de X son 0, 1 y 2). Calcular las probabilidades $P(X = 0)$ y $P(X = 1)$.
   c) Representar la función de distribución F_X cuando $p = \frac{1}{3}$.
8. Dada una variable aleatoria X cuyos valores posibles son $1, 2, \ldots, n$. Admitamos que la probabilidad $P(X = k)$ es proporcional a k. Determinar la constante de proporcionalidad, la función de masa de probabilidad p_X, y la de distribución F_X.
9. Dada una variable aleatoria X cuyos valores posibles son $0, 1, 2, 3, \ldots$. Admitamos que $P(X = k)$ es proporcional a $c^k/k!$, donde c es un número real fijo. Determinar la constante de proporcionalidad y la función de masa de probabilidad p.
10. a) Se lanza un dado correcto. El espacio muestral es $S = \{1, 2, 3, 4, 5, 6\}$. Si el número de puntos conseguidos es impar, un jugador recibe un dólar; en el caso contrario debe pagar un dólar. Sea X la variable aleatoria que mide el resultado (número de dólares) en cada partida del juego. (Los valores posibles de X son $+1$ y -1). Describir y representar las funciones de masa de probabilidad p_X y de distribución F_X.
    b) Se lanza una moneda perfecta. El espacio muestral es $S = \{H, T\}$ (cara o cruz). Si el resultado es cara, un jugador recibe un dólar; si es cruz paga un dólar. Sea Y la variable aleatoria que mide el resultado (número de dólares) en cada partida del juego. Demostrar que las funciones de masa de probabilidad p_Y y de distribución F_Y son idénticas a las de la parte a) de este ejercicio. Este ejemplo pone de manifiesto que variables aleatorias distintas pueden tener la misma función de distribución dada F. (¿Por qué?) Se dice que tales variables aleatorias están *distribuidas idénticamente*. Cada teorema relativo a una función de distribución particular es aplicable a cualquiera de las infinitas variables aleatorias que tengan esta distribución.
11. Se sabe que el número de minutos que uno tiene que esperar un tren en cierta estación del metro es una variable aleatoria X con la siguiente función de masa de probabilidad:

$$p(t) = 0 \quad \text{salvo para } t = 3k/10 \qquad \text{si} \qquad k = 0, 1, 2, \ldots, 10.$$

$$p(t) = \tfrac{1}{12} \qquad \text{si} \quad t = 0, 0.3, 0.6, 0.9, 2.1, 2.4, 2.7, 3.0.$$

$$p(t) = \tfrac{1}{9} \qquad \text{si} \quad t = 1.2, 1.5, 1.8.$$

    Representar la correspondiente función de distribución F. Sea A el suceso consistente en que uno tiene que esperar entre 0 y 2 minutos (incluyendo 0 y 2), y sea B el suceso de tener que esperar entre 1 y 3 minutos (incluyendo 1 y 3). Calcular las probabilidades: $P(A)$, $P(B)$, $P(A \cap B)$, $P(B A)$, $P(A \cup B)$.
12. a) Si $0 < p < 1$ y $q = 1 - p$, demostrar que

$$\binom{n}{k} p^k q^{n-k} = \frac{(np)^k}{k!}\left(1 - \frac{np}{n}\right)^n Q_n,$$

siendo

$$Q_n = \frac{\prod_{r=2}^{k}\left(1 - \frac{r-1}{n}\right)}{(1-p)^k} .$$

b) Dado $\lambda > 0$, sea $p = \lambda/n$ para $n > \lambda$. Demostrar que $Q_n \to 1$ cuando $n \to \infty$ y que

$$\binom{n}{k} p^k q^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda} \text{ cuando } n \to \infty .$$

Este resultado sugiere que para n grande e y pequeño, la distribución binomial es aproximadamente la misma que la de Poisson, con tal que el producto np sea casi constante; esta constante es el parámetro λ de la distribución de Poisson.

14.9 Distribuciones continuas. Funciones de densidad

Sea X una variable aleatoria unidimensional y F su función de distribución, de modo que $F(t) = P(X \leq t)$ para todo t real. Si la probabilidad $P(X = t)$ es cero para todo t, entonces, en virtud del teorema 14.5, F es continua en todo el eje real. En este caso F se llama *distribución continua* y X es una *variable aleatoria continua.* Si existe la derivada F' y es continua en el intervalo $[a, t]$ puede hacerse uso del segundo teorema fundamental del cálculo y escribir:

(14.13)
$$F(t) - F(a) = \int_a^t f(u)\, du ,$$

siendo f la derivada de F. La diferencia $F(t) - F(a)$ es, naturalmente, la probabilidad $P(a < X \leq t)$, y la desigualdad (14.13) expresa esta probabilidad como una integral.

En algunas ocasiones la función de distribución F puede expresarse como una integral de la forma (14.13), en la que el integrando f es integrable pero no necesariamente continua. Siempre que una igualdad como la (14.13) es válida para todos los intervalos $[a, t]$, el integrando f se llama *función de densidad de probabilidad* de la variable aleatoria X (o de la distribución F) con tal que f sea no negativa. Dicho de otro modo, tenemos la siguiente definición:

DEFINICIÓN DE FUNCIÓN DE DENSIDAD DE PROBABILIDAD. *Sea X una variable aleatoria continua con función de distribución F. Una función f no negativa se denomina densidad de probabilidad de X (o de F) si es integrable en todo el intervalo $[a, t]$ y si*

(14.14)
$$F(t) - F(a) = \int_a^t f(u)\, du .$$

Si en (14.14) hacemos que $a \to -\infty$, entonces $F(a) \to 0$ y obtenemos la importante fórmula

$$(14.15) \qquad F(t) = P(X \leq t) = \int_{-\infty}^{t} f(u)\, du\,,$$

válida para todo t real. Si hacemos ahora $t \to +\infty$ y recordamos que $F(t) \to 1$ encontramos que

$$(14.16) \qquad \int_{-\infty}^{+\infty} f(u)\, du = 1\,.$$

Para las variables aleatorias discretas la suma de todas las probabilidades $P(X = t)$ es igual a 1. La fórmula (14.16) es la versión de la misma propiedad adaptada a variables aleatorias continuas. También existe una estrecha analogía entre las fórmulas (14.11) y (14.15). La función densidad f desempeña, para las distribuciones continuas, el mismo papel que la función de masa de probabilidad p para distribuciones discretas; la integración reemplaza la suma en el cálculo de las probabilidades. Existe una diferencia importante, sin embargo. En el caso discreto $p(t)$ es la probabilidad de que $X = t$, pero en el continuo $f(t)$ *no es* la probabilidad de que $(X = t)$. En efecto, esta probabilidad es cero debido a que F es continua para todo t, lo que equivale a que para una distribución continua es:

$$P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X < b)\,.$$

Si F tiene una densidad f cada una de esas probabilidades es igual a la integral $\int_a^b f(u)\, du$.

Observación: Una distribución dada puede tener más de una densidad puesto que el valor del integrando de (14.14) puede modificarse en un número finito de puntos sin que por ello varíe la integral. Pero si f es *continua* en t entonces $f(t) = F'(t)$; en este caso el valor de la función densidad en t está determinado con unicidad por F.

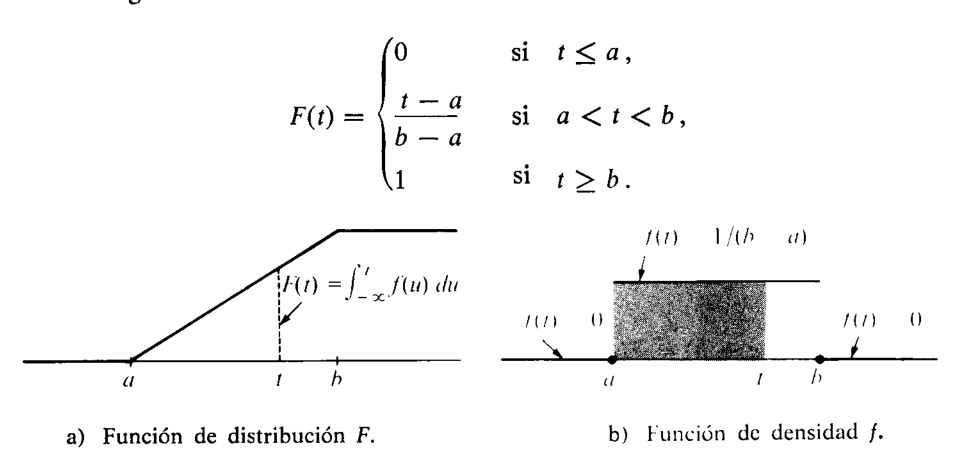
Como f es no negativa, se puede interpretar geométricamente la ecuación (14.14) como el área del conjunto de ordenadas de f situadas a la izquierda de la recta $x = t$. El área de todo el conjunto de ordenadas es igual a 1. El área de la parte del conjunto de ordenadas situada encima de un intervalo dado (sea abierto, cerrado o semiabierto) es la probabilidad de que la variable aleatoria X tome un valor en tal intervalo. La figura 14.6 muestra un ejemplo de una función F de distribución continua y de su función de densidad f. La ordenada $F(t)$ en la figura 14.6 a) es igual al área de la región sombreada en la figura 14.6 b).

En las secciones que siguen vamos a ver algunos ejemplos importantes de distribuciones continuas.

14.10 Distribución uniforme sobre un intervalo

Se dice que una variable aleatoria uni-dimensional X tiene una función de distribución uniforme F sobre un intervalo finito $[a, b]$ si F viene dada por las fórmulas siguientes:

$$F(t) = \begin{cases} 0 & \text{si} \quad t \leq a, \\ \dfrac{t-a}{b-a} & \text{si} \quad a < t < b, \\ 1 & \text{si} \quad t \geq b. \end{cases}$$



a) Función de distribución F.

b) Función de densidad f.

FIGURA 14.6 *Una función de distribución uniforme sobre un intervalo* $[a, b]$ *y la correspondiente función de densidad.*

Ésta es una distribución continua cuya gráfica es la de la figura 14.6 a).

La derivada $F'(t)$ existe para todo valor de t excepto en los puntos a y b, y podemos escribir

$$F(t) = \int_{-\infty}^{t} f(u)\, du,$$

en donde f es la función de densidad, definida como sigue:

$$f(t) = \begin{cases} 1/(b-a) & \text{si} \quad a < t < b, \\ 0 & \text{para los demás valores de } t \end{cases}$$

La gráfica de f está representada en la figura 14.6 b).

El teorema que sigue caracteriza las funciones de distribución de otro modo.

TEOREMA 14.7. *Sea X una variable aleatoria uni-dimensional con todos sus valores en un intervalo finito $[a, b]$, y sea F la función de distribución de X. La*

función F es uniforme sobre $[a, b]$ si y sólo si

$$P(X \in I) = P(X \in J) \tag{14.17}$$

para todo par de subintervalos I y J de $[a, b]$ que tengan la misma longitud, en cuyo caso tenemos

$$P(X \in I) = \frac{h}{b-a},$$

en donde h es la longitud de I.

Demostración. Supongamos primero que X tiene una función de distribución sobre $[a, b]$. Si $[c, c+h]$ es un subintervalo cualquiera de $[a, b]$ de longitud h tenemos

$$P(c \leq X \leq c+h) = F(c+h) - F(c) = \frac{c+h-a}{b-a} - \frac{c-a}{b-a} = \frac{h}{b-a}.$$

Esto demuestra que $P(X \in I) = P(X \in J) = h/(b-a)$ para todo par de subintervalos I y J de $[a, b]$ de longitud h.

Para demostrar el recíproco, supongamos que X satisface (14.17). Observemos primero que $F(t) = 0$ si $t < a$ y $F(t) = 1$ si $t > b$, ya que X tiene todos sus valores en $[a, b]$,

Introduzcamos una nueva función g definida en el intervalo semi abierto $(0, b-a]$ mediante la ecuación

$$g(u) = P(a < X \leq a+u) \quad \text{si} \quad 0 < u \leq b-a.$$

Aplicando la aditividad y la propiedad (14.17) encontramos

$$\begin{aligned} g(u+v) &= P(a < X \leq a+u+v) \\ &= P(a < X \leq a+u) + P(a+u < X \leq a+u+v) \\ &= g(u) + P(a < X \leq a+v) = g(u) + g(v), \end{aligned}$$

con tal que $0 < u + v \leq b - a$. Esto es, g satisface la ecuación funcional

$$g(u+v) = g(u) + g(v) \tag{14.18}$$

para todas las funciones u y v tales que $u > 0$, $v > 0$, $u + v \leq b - a$. Esta es la llamada *ecuación funcional de Cauchy*. Demostraremos en breve que toda solu-

ción no negativa de la ecuación funcional de Cauchy viene dada por

$$g(u) = \frac{u}{b-a} g(b-a) \quad \text{para} \quad 0 < u \leq b-a .$$

Aplicando esto en la ecuación (14.18) encontramos que para $0 < u \leq b - a$ tenemos

$$P(a < X \leq a+u) = \frac{u}{b-a} P(a < X \leq b) = \frac{u}{b-a}$$

ya que $P(a < X \leq b) = 1$. Es decir,

$$F(a+u) - F(a) = \frac{u}{b-a} \quad \text{si} \quad 0 < u \leq b-a .$$

Pongamos $t = a + u$ y esta última igualdad se puede escribir en la forma

$$F(t) - F(a) = \frac{t-a}{b-a} \quad \text{si} \quad a < t \leq b .$$

Pero $F(a) = 0$ ya que F es continua por la derecha. Luego

$$F(t) = \frac{t-a}{b-a} \quad \text{si} \quad a \leq t \leq b ,$$

lo que demuestra que F es uniforme en $[a, b]$.

TEOREMA 14.8. SOLUCIÓN DE LA ECUACIÓN FUNCIONAL DE CAUCHY. *Si g es una función de valores reales definida en un intervalo semiabierto $(0, c]$ y que satisface las propiedades siguientes*

a) $g(u+v) = g(u) + g(v)$ *siempre que u, v y $u+v$ pertenezcan a* $(0, c]$.

b) *g es no negativa en* $(0, c]$.

Entonces g viene dada por la fórmula

$$g(u) = \frac{u}{c} g(c) \quad \textit{para} \quad 0 < u \leq c .$$

Demostración. Introduciendo un cambio de escala podemos reducir la demostración al caso particular en el que $c = 1$. En efecto, pongamos

$$G(x) = g(cx) \quad \text{para} \quad 0 < x \leq 1 .$$

Entonces G es no negativa y satisface la ecuación funcional de Cauchy

$$G(x + y) = G(x) + G(y)$$

siempre que x, y, y $x + y$ pertenezcan a $(0, 1]$. Si demostramos que

(14.19) $$G(x) = xG(1) \quad \text{para} \quad 0 < x \leq 1$$

resulta que $g(cx) = xg(c)$, o que $g(u) = (u/c)g(c)$ para $0 < u \leq c$.

Si x está en $(0, 1]$ entonces $x/2$ también pertenece a $(0, 1]$ y tenemos

$$G(x) = G\left(\frac{x}{2}\right) + G\left(\frac{x}{2}\right) = 2G\left(\frac{x}{2}\right).$$

Por inducción, para cada x de $(0,1]$ tenemos

(14.20) $$G(x) = nG\left(\frac{x}{n}\right) \quad \text{para} \quad n = 1, 2, 3, \ldots .$$

Del mismo modo, si y y my están en $(0, 1]$ tenemos

$$G(my) = mG(y) \quad \text{para} \quad m = 1, 2, 3, \ldots .$$

Haciendo $y = x/n$ y aplicando (14.20) obtenemos

$$G\left(\frac{m}{n}x\right) = \frac{m}{n}G(x)$$

si x y mx/n están en $(0, 1]$. Dicho de otro modo, tenemos

(14.21) $$G(rx) = rG(x)$$

para todo número racional positivo r tal que x y rx estén en $(0, 1]$.

Tomemos ahora un x cualquiera en el intervalo abierto $(0, 1)$ y sean $\{r_n\}$ y $\{R_n\}$ de dos sucesiones números| racionales del intervalo $(0, 1]$ tales que

$$r_n < x < R_n \qquad \text{y que} \qquad \lim_{n\to\infty} r_n = \lim_{n\to\infty} R_n = x .$$

La ecuación funcional de Cauchy y la propiedad de ser G no negativa demuestran

que $G(x + y) \geq G(x)$ así que G es monótona creciente en $(0, 1]$. Por consiguiente

$$G(r_n) \leq G(x) \leq G(R_n).$$

Aplicando (14.21) podemos poner esto en la forma

$$r_n G(1) \leq G(x) \leq R_n G(1).$$

Haciendo que $n \to \infty$ encontramos $xG(1) \leq G(x) \leq x\,G(1)$, con lo que $G(x) = xG(1)$, lo cual demuestra (14.19).

Observación: Las distribuciones uniformes se utilizan a menudo en experimentos cuyos resultados son puntos seleccionados al azar en un intervalo $[a, b]$, o en experimentos que se refieren a un intervalo $[a, b]$ como un blanco, en el que es imposible hacer puntería. Los términos «al azar» e «imposible hacer puntería» se interpretan ordinariamente para significar que si I es un subintervalo cualquiera de $[a, b]$ entonces la probabilidad $P(X \in I)$ depende solamente de la longitud de I y no de su posición en $[a, b]$. El teorema 14.7 demuestra que las distribuciones uniformes son las únicas distribuciones que tienen esa propiedad.

Volvamos ahora a las cuestiones de probabilidad formuladas al comienzo de este capítulo.

EJEMPLO. Un segmento rectilíneo se descompone en dos partes, eligiendo el punto de subdivisión al azar. Designemos con X la variable aleatoria que mide la razón de la longitud de la parte izquierda a la de la parte derecha. Determinar la función de distribución de probabilidad F_X.

Solución. Utilicemos el intervalo $[0, 1]$ para representar el segmento rectilíneo y expresemos el punto de subdivisión mediante la variable aleatoria $Y(\omega) = \omega$ para cada ω en $(0, 1)$. Ya que el punto de subdivisión se elige al azar supongamos que Y tiene una función de distribución uniforme F_Y sobre $[0, 1]$. Luego

$$F_Y(t) = t \qquad \text{para } 0 \leq t \leq 1.$$

Si el segmento se parte en ω, entonces $\omega/(1 - \omega)$ es la razón de la longitud de la parte izquierda a la de la parte derecha. Por tanto $X(\omega) = \omega/(1 - \omega)$.

Si $t < 0$ tenemos $F_X(t) = 0$ ya que la razón $X(\omega)$ no puede ser negativa. Si $t \geq 0$, la desigualdad $X(\omega) \leq t$ es equivalente a $\omega/(1 - \omega) \leq t$, la cual equivale a $\omega \leq t/(1 + t)$. Por consiguiente

$$F_X(t) = P(X \leq t) = P\left(Y \leq \frac{t}{1 + t}\right) = F_Y\left(\frac{t}{1 + t}\right) = \frac{t}{1 + t}$$

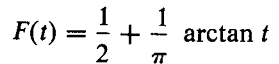
puesto que $0 \leq t/(1 + t) < 1$.

Ahora podemos calcular diversas probabilidades. Por ejemplo, la probabilidad de que las dos partes tengan la misma longitud es $P(X = 1) = 0$. En efecto, puesto que F_X es una distribución continua, la probabilidad de que X tome un valor particular cualquiera es cero.

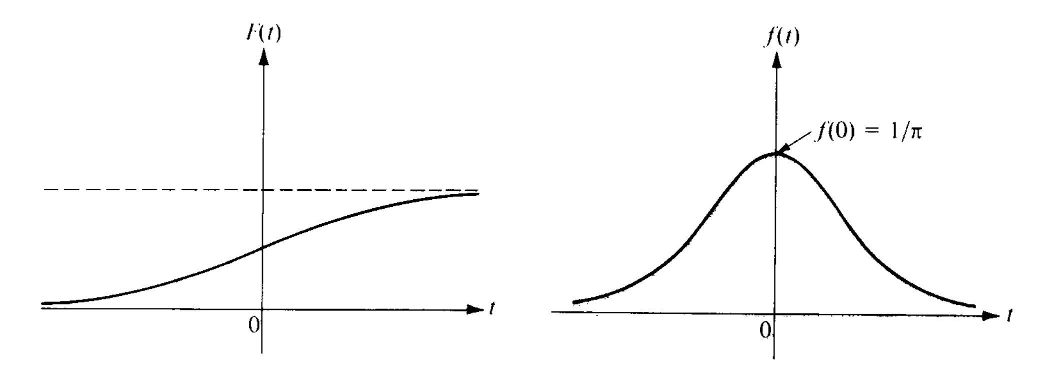
La probabilidad de que el segmento de la izquierda sea por lo menos de longitud doble que el de la derecha es $P(X \geq 2) = 1 - P(X < 2) = 1 - \frac{2}{3} = \frac{1}{3}$. Análogamente, la probabilidad de que el segmento de la derecha sea por lo menos de longitud doble que el de la izquierda es $P(X \leq \frac{1}{2}) = \frac{1}{3}$. La probabilidad de que la longitud del segmento total sea por lo menos de longitud doble que la de la parte menor es $P(X \geq 2) + P(X \leq \frac{1}{2}) = \frac{2}{3}$.

14.11 Distribución de Cauchy

Se dice que una variable aleatoria X tiene una distribución de Cauchy F si



$$F(t) = \frac{1}{2} + \frac{1}{\pi} \arctan t$$



a) Función de distribución F.

b) Función de densidad f.

FIGURA 14.7 *Función de distribución de Cauchy y la correspondiente función de densidad.*

para todo t real. Esta función tiene derivada continua para todo valor de t; la correspondiente función de densidad f viene dada por la fórmula

$$f(t) = \frac{1}{\pi(1 + t^2)}.$$

Las gráficas de f y F son las de las figuras 14.7 b) y a), respectivamente.

El experimento siguiente nos conduce a una distribución de Cauchy. A una agu-

ja que puede girar alrededor del punto $(-1, 0)$ del eje x se la impulsa para que gire y se la deja parar espontáneamente. Un resultado del experimento es el ángulo θ que forma la aguja con el eje x; θ se mide de modo que $-\frac{1}{2}\pi < \theta \leq \frac{1}{2}\pi$. Razonando de manera parecida a como se ha hecho en el ejemplo 1 se demuestra que la variable aleatoria X definida por $X(\theta) = \theta$ tiene una distribución uniforme sobre el intervalo $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$, con densidad $f_X(t) = \frac{1}{\pi}$ si $-\frac{1}{2}\pi < t < \frac{1}{2}\pi$. Consideremos ahora la variable aleatoria Y definida por

$$Y(\theta) = \tan\theta.$$

Demostraremos ahora que Y tiene una distribución de Cauchy F_Y si X tiene distribución uniforme sobre $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$.

Si $a < t$ sea $\alpha = \operatorname{arc\,tan} a$ y $\theta = \operatorname{arc\,tan} t$. Tenemos entonces,

$$F_Y(t) - F_Y(a) = P(a < Y \leq t) = P(\alpha < X \leq \theta) = \int_\alpha^\theta f_X(u)\,du = \frac{\theta - \alpha}{\pi}.$$

Ya que $\alpha \to -\pi/2$ cuando $a \to -\infty$ encontramos:

$$F_Y(t) = \frac{\theta + \frac{1}{2}\pi}{\pi} = \frac{1}{\pi}\arctan t + \frac{1}{2}.$$

Esto nos demuestra que Y tiene una distribución de Cauchy.

14.12 Ejercicios

1. Una variable aleatoria X tiene función de distribución continua F, siendo

$$F(t) = \begin{cases} 0 & \text{si } t \leq 0, \\ ct & \text{si } 0 \leq t \leq 1, \\ 1 & \text{si } t > 1. \end{cases}$$

   a) Determinar la constante c y describir la función densidad f.
   b) Calcular las probabilidades $P(X = \frac{1}{3})$, $P(X < \frac{1}{3})$, $P(|X| < \frac{1}{3})$.

2. Sea $f(t) = c\,|\operatorname{sen} t|$ para $|t| < \pi/2$ y $f(t) = 0$ para los demás valores de t. Determinar el valor de la constante c de manera que f sea la densidad de una función de distribución continua F. Describir, también, F y representarla.
3. Resolver el ejercicio 2 si $f(t) = c(4t - 2t^2)$ para $0 \leq t \leq 2$, y $f(t) = 0$ para los demás valores de t.
4. El tiempo en minutos que una persona espera un autobús es una variable aleatoria con

función de densidad f dada por las fórmulas siguientes:

$$f(t) = \tfrac{1}{2} \text{ para } 0 < t < 1,\ f(t) = \tfrac{1}{4} \text{ para } 2 < t < 4,\ f(t) = 0 \text{ para los demás valores de } t.$$

Calcular la probabilidad de que el tiempo de espera sea a) mayor que un minuto; b) mayor que dos minutos; c) mayor que tres minutos.

5. Una variable aleatoria X tiene una función de distribución continua F y una densidad de probabilidad f. La densidad tiene las propiedades siguientes: $f(t) = 0$ si $t < \frac{1}{4}$, $f(\frac{1}{4}) = 1$, $f(t)$ es lineal si $\frac{1}{4} \le t \le \frac{1}{2}$, $f(1-t) = f(t)$ para todo t.
   a) Representar f.
   b) Dar un conjunto de fórmulas que determinan F y representarla.
   c) Calcular las probabilidades: $P(X<1)$, $P(X<\frac{3}{4})$, $P(X<\frac{1}{2})$, $P(X\le\frac{1}{4})$, $P(\frac{1}{2}<X<\frac{5}{8})$.
6. Una variable aleatoria X tiene una distribución uniforme sobre $[-3,3]$.
   a) Calcular $P(X = 2)$, $P(X < 2)$, $P(|X| < 2)$, $P(|X-2| < 2)$.
   b) Hallar un valor de t para el que $P(X > t) = \frac{1}{3}$.
7. La Compañía Lethe Subway tiene señalado un tren en dirección norte cada 30 minutos en cierta estación. Sea X la variable aleatoria que cuenta el número de minutos que un hombre, que entró en la estación en un tiempo aleatorio, debe esperar el próximo tren. Supongamos que X tiene una distribución uniforme en el intervalo $[0, 30]$. (Este es el sentido que debemos atribuir a la frase, el señor entra en la estación en un «tiempo aleatorio».)
   a) Para cada $k = 5, 10, 15, 20, 25, 30$, calcular la probabilidad que tiene de esperar por lo menos k minutos para el próximo tren.
   b) Una compañía competidora, la Styx Subway, está autorizada para anunciar un tren cada 30 minutos en dirección norte, en la misma estación, pero las llegadas de dos trenes de distinta compañía deben diferir por lo menos 5 minutos. Supongamos que los viajeros entran en la estación en tiempos aleatorios y siempre toman el primer tren que llega. Demostrar que la compañía Styx puede combinar su horario de manera que recoja hasta cinco veces el número de viajeros que su competidora.
8. Sea X una variable aleatoria con una distribución uniforme F_X sobre el intervalo $[0, 1]$. Pongamos $Y = aX + b$, donde $a > 0$. Determinar la función de distribución F_Y y dibujar su gráfica.
9. Una ruleta recorre los enteros de 0 a 36, distribuidos en 37 arcos iguales. Se pone en marcha la ruleta y se la deja parar libremente, y se anota el punto de la circunferencia junto al cual se para. Considérese este punto de parada como una variable aleatoria X con una distribución uniforme. Calcular la probabilidad de que X esté en un arco que contenga a) el entero 0; b) un entero n en el intervalo $11 \le n \le 20$; c) un entero impar.
10. Se dice que una variable aleatoria tiene una distribución de Cauchy con parámetros a y b, siendo $a > 0$, si su función de densidad viene dada por

$$f(t) = \frac{1}{\pi}\frac{a}{a^2 + (t-b)^2}.$$

Comprobar que la integral de f entre $-\infty$ y $+\infty$ es 1, y determinar la función de distribución F.

11. Sea $f_1(t) = 1$ para $0 < t < 1$, y $f_1(t) = 0$ para los demás valores de t. Definamos una sucesión funcional $\{f_n\}$ mediante la fórmula recurrente

$$f_{n+1}(x) = \int_{-\infty}^{\infty} f_1(x-t) f_n(t)\, dt.$$

a) Demostrar que $f_{n+1}(x) = \int_{x-1}^{x} f_n(t)\,dt$.
b) Esbozar las gráficas de f_1, f_2 y f_3.

12. Respecto al ejercicio 11. Demostrar que cada función f_n es una densidad de probabilidad.

14.13 Distribuciones exponenciales

Sea λ una constante positiva. Se dice que una variable aleatoria uni-dimensional X tiene una función de distribución exponencial F con parámetro λ si

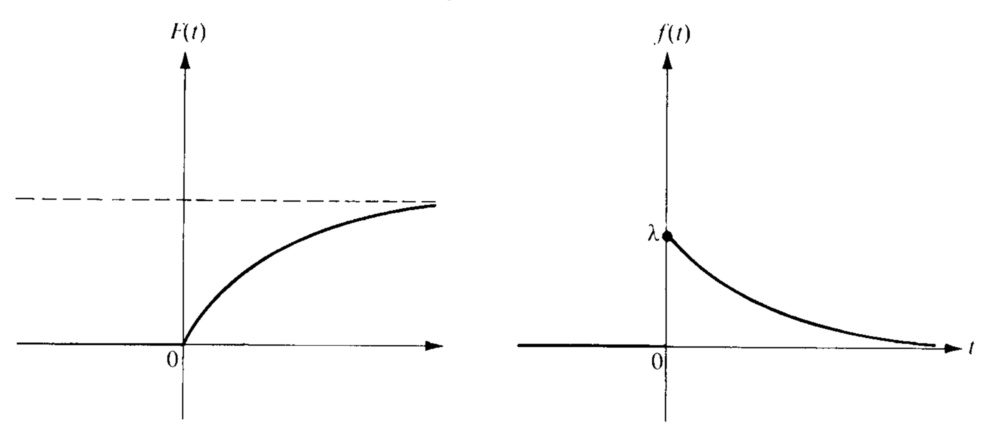
$$F(t) = \begin{cases} 1 - e^{-\lambda t} & \text{para } t \geq 0, \\ 0 & \text{para } t < 0. \end{cases}$$

La función densidad f correspondiente viene dada por las fórmulas

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{para } t \geq 0, \\ 0 & \text{para } t < 0. \end{cases}$$

Las gráficas de F y f son parecidas a las de la figura 14.8.

Las distribuciones exponenciales tienen una propiedad característica que sugiere su aplicación en ciertos problemas relativos a desintegración radiactiva, accidentes de tráfico, y averías en aparatos de equipos electrónicos tales como tubos de vacío. Esta propiedad es análoga a la que caracteriza las distribuciones uniformes y puede describirse del siguiente modo.



a) Función de distribución F.

b) Función de densidad f.

FIGRA 14.8 *Distribución exponencial y función de densidad correspondiente.*

Designemos con X el tiempo transcurrido hasta que una pieza del equipo falla, y sea F la función de distribución de X. Supongamos que $F(t) = 0$ para $t \leq 0$, y por el momento no imponemos otras restricciones a F. Si $t > 0$, entonces $X \leq t$ es el suceso «el fallo sucede en el intervalo $[0, t]$». Luego $X > t$ es el suceso complementario, «ningún fallo ocurre en el intervalo $[0, t]$».

Supongamos que no se presenta fallo en el intervalo $[0, t]$. ¿Cuál es la probabilidad de supervivencia, es decir de que no haya fallo en el intervalo $[t, t + s]$? Esta es una cuestión de probabilidades condicionadas. Queremos determinar $P(X > t + s | X > t)$, la probabilidad condicionada de que no exista fallo en el intervalo $[0, t + s]$, dado que no haya existido en el intervalo $[0, t]$.

Según la definición de probabilidad condicionada tenemos

$$P(X > t + s \mid X > t) = \frac{P[(X > t + s) \cap (X > t)]}{P(X > t)} = \frac{P(X > t + s)}{P(X > t)}. \tag{14.22}$$

Supongamos ahora que F es una distribución exponencial con parámetro $\lambda > 0$. Entonces $F(t) = 1 - e^{-\lambda t}$ para $t > 0$, y $P(X > t) = 1 - P(X \leq t) = e^{-\lambda t}$. Luego la ecuación (14.22) se transforma en

$$P(X > t + s \mid X > t) = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s).$$

Dicho de otro modo, si el órgano o pieza del equipo sobrevive en el intervalo $[0, t]$, la probabilidad de que continúe sobreviviendo en el intervalo $[t, t + s]$ es igual a la probabilidad de supervivencia en el intervalo $[0, s]$ de la misma longitud. Esto es, la probabilidad de supervivencia depende tan sólo de la amplitud del intervalo de tiempo y no de la edad del aparato. Expresado con la función de distribución F, esta propiedad establece que

$$\frac{1 - F(t + s)}{1 - F(t)} = 1 - F(s) \quad \text{para todo } t > 0 \text{ y todo } s > 0. \tag{14.23}$$

El teorema que sigue demuestra que las distribuciones exponenciales son las únicas distribuciones de probabilidad con esa propiedad.

TEOREMA 14.9. *Si F es una función de distribución de probabilidad que satisface la ecuación funcional* (14.23), *siendo $F(t) < 1$ para $t > 0$, existe entonces una constante positiva $\lambda > 0$ tal que*

$$F(t) = 1 - e^{-\lambda t} \quad \textit{para todo } t > 0.$$

Demostración. Sea $g(t) = -\log [1 - F(t)]$ para $t > 0$. Entonces $1 - F(t) = e^{-g(t)}$, así que para demostrar el teorema es suficiente probar que $g(t) = \lambda t$ para un cierto $\lambda > 0$.

Ahora bien g es no negativa y satisface la ecuación funcional de Cauchy

$$g(t + s) = g(t) + g(s)$$

para todo $t > 0$ y todo $s > 0$. Por consiguiente, aplicando el teorema 14.8 con $c = 1$, deducimos que $g(t) = tg(1)$ para $0 < t \leq 1$. Pongamos $\lambda = g(1)$. Entonces $\lambda = -\log [1 - F(1)] > 0$, y por tanto $g(t) = \lambda t$ para $0 < t \leq 1$.

Para demostrar que $g(t) = \lambda t$ para todo $t > 0$, pongamos $G(t) = g(t) - \lambda t$. La función G también satisface la ecuación funcional de Cauchy. Además, G es periódica con período 1 debido a que $G(t + 1) = G(t) + G(1)$ y $G(1) = 0$. Puesto que G es idénticamente 0 en $(0, 1]$ la periodicidad demuestra que $G(t) = 0$, para todo $t > 0$. Es decir, $g(t) = \lambda t$ para todo $t > 0$, lo cual completa la demostración.

EJEMPLO 1. Sea X una variable aleatoria que mide el tiempo de vida (en horas) de un cierto tipo de tubo de vacío. Supongamos que X tenga una distribución exponencial con parámetro $\lambda = 0{,}001$. El fabricante desea garantizar esos tubos para una duración de T horas. Determinar T de modo que $P(X > T) = 0{,}95$.

Solución. La función de distribución viene dada por $F(t) = 1 - e^{-\lambda t}$ para $t > 0$, en donde $\lambda = 0{,}001$. Puesto que $P(X > T) = 1 - F(T) = e^{-\lambda T}$, elegimos T para que $e^{-\lambda T} = 0{,}95$. Luego $T = -(\log 0{,}95)/\lambda = -1\,000 \log 0{,}95 = 51{,}25+$.

EJEMPLO 2. Consideremos la variable aleatoria del ejemplo 1, pero con un valor de λ no especificado. El razonamiento que sigue sugiere un procedimiento razonable para determinar λ. Empezamos con un número inicial de tubos de vacío en el tiempo $t = 0$, y designemos con $g(t)$ el número de tubos que aún funcionan al cabo de t horas. La razón $[g(0) - g(t)]/g(0)$ es la fracción del número original que ha fallado en el tiempo t. Puesto que la probabilidad de que un tubo falle en el tiempo t es $1 - e^{-\lambda t}$, parece razonable esperar que la ecuación

$$\frac{g(0) - g(t)}{g(0)} = 1 - e^{-\lambda t} \tag{14.24}$$

sea una buena aproximación de la realidad. Si suponemos (14.24) obtenemos

$$g(t) = g(0)e^{-\lambda t}.$$

Es decir, en las hipótesis (14.24), el número $g(t)$ obedece a una ley de deterioro

exponencial con deterioro constante λ. El deterioro constante puede calcularse en función de la vida media. Si t_1 es la vida media entonces $\frac{1}{2} = g(t_1)/g(0) = e^{-\lambda t_1}$, así que $\lambda = (\log 2)/t_1$. Por ejemplo, si se sabe que la vida media de una gran muestra de tubos es 693 horas, obtenemos $\lambda = (\log 2)/693 = 0{,}001$.

14.14 Distribuciones normales

Sean m y σ dos números reales fijos, siendo $\sigma > 0$. Se dice que una variable aleatoria X tiene *distribución normal* de media m y varianza σ^2 si la función de densidad es

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-[(t-m)/\sigma]^2/2}$$

para todo t real. La función de distribución correspondiente F es, naturalmente, la integral

$$F(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-[(u-m)/\sigma]^2/2}\, du\,.$$

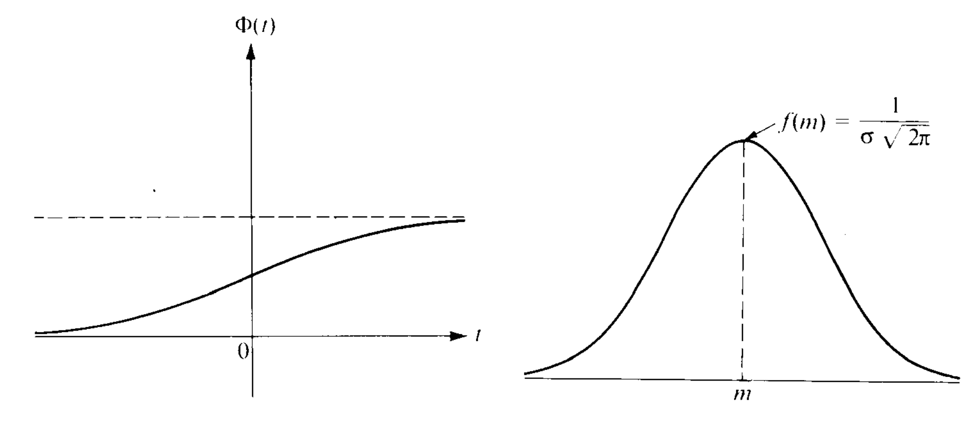


FIGURA 14.9 *Función de distribución normal canónica:* m = 0, $\sigma = 1$.

FIGURA 14.10 *Función densidad de una distribución normal de media m y varianza* σ^2.

Es evidente que esta función F es monótona creciente y continua en todo el eje real, y que tiende a cero cuando $t \to -\infty$. También se puede demostrar que $F(t) \to 1$ cuando $t \to +\infty$. (Véase ejercicio 7 de la sección 14.16.)

El caso particular $m = 0$, $\sigma = 1$ es la distribución normal canónica. En tal caso la función F se representa ordinariamente con la letra Φ. Así pues,

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2}\, du\,.$$

TABLA 14.1 *Valores de la función de distribución normal canónica.*

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2}\, du.$$

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0.9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

El caso general puede reducirse al canónico introduciendo el cambio de variable $v = (u - m)/\sigma$ en la integral que define F. Esto nos lleva a la fórmula

$$F(t) = \Phi\left(\frac{t - m}{\sigma}\right).$$

La tabla 14.1 es una tabla de valores de $\Phi(t)$ con cuatro cifras decimales para valores de t espaciados en intervalos de amplitud 0,01 desde $t = 0{,}00$ a $t = 3{,}69$. La gráfica de $\Phi(t)$ está representada en la figura 14.9. La gráfica de la densidad f es la famosa «curva de campana», dibujada en la figura 14.10. La parte alta de la campana corresponde a la media m. Para valores grandes de σ la curva se achata, y para valores pequeños de σ la curva de campana se estrecha y en el máximo tiene un aspecto más pronunciado, como en la figura 14.10.

La distribución normal es una de las más importantes entre las distribuciones continuas. Muchas variables aleatorias que se presentan en la naturaleza se comportan como si su distribución fuera normal o aproximadamente normal. Como ejemplos citaremos la medida de la estatura de los individuos de una gran población, ciertas medidas sobre grandes poblaciones de organismos vivos que se presentan en Biología, y los errores de observación encontrados al hacer un gran número de mediciones. En Física, la ley de Maxwell de las velocidades implica que la función de distribución de la velocidad, en cualquier dirección dada, de una molécula de masa M en un gas a la temperatura absoluta T es normal con media 0 y varianza $M/(kT)$ donde k es una constante (constante de Boltzmann).

La distribución normal es también de importancia teórica debido a que puede utilizarse para aproximar las distribuciones de muchos fenómenos aleatorios. Un ejemplo es la distribución binomial con parámetros n y p. Si X es una variable aleatoria con distribución binomial de parámetros n y p, la probabilidad $P(a \leq X \leq b)$ viene dada por la suma

$$\sum_{k=a}^{b} \binom{n}{k} p^k q^{n-k},$$

donde $q = 1 - p$. Para n grande son necesarios cálculos laboriosos para hallar esta suma. En la práctica se evitan estos cálculos mediante la fórmula aproximada

$$\sum_{k=a}^{b} \binom{n}{k} p^k q^{n-k} \sim \Phi\left(\frac{b - np + \frac{1}{2}}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np - \frac{1}{2}}{\sqrt{npq}}\right), \tag{14.25}$$

en donde el símbolo $\sim$ significa que los dos miembros de 14.25 son asintóticamente iguales; esto es, que la razón del primer miembro al segundo tiende a 1 cuando $n \to \infty$. El límite expresado en (14.25) es una caso particular del llamado

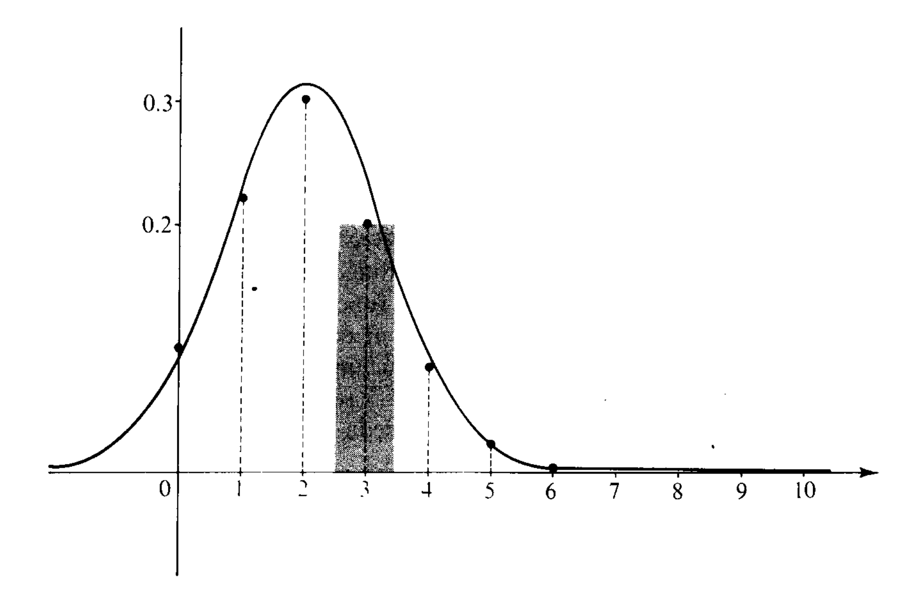


FIGURA 14.11 *Función densidad de una distribución normal considerada como aproximación de la función de masa de probabilidad de una distribución binomial.*

teorema central del límite del Cálculo de probabilidades. Este teorema (discutido con detalle en la sección 14.30) explica la importancia teórica de la distribución normal.

La figura 14.11 da una idea de la fórmula de aproximación (14.25) y muestra que puede ser exacta para un valor relativamente pequeño de n. Las rectas punteadas son las ordenadas de la función de masa de probabilidad p de una distribución binomial de parámetros $n = 10$ y $p = 1/5$. Esas ordenadas fueron calculadas con la fórmula

$$p(t) = P(X = t) = \binom{10}{t}\left(\frac{1}{5}\right)^t \left(\frac{4}{5}\right)^{10-t} \quad \text{para} \quad t = 0, 1, 2, \ldots, 10.$$

Las ordenadas para $t =$ 7, 8, 9, y 10 no han sido dibujadas porque sus valores numéricos son muy próximos a cero. Por ejemplo, $p(10) = (1/5)^{10} = 2^{10}/10^{10} =$ $= 0{,}0000001024$. La curva dibujada es la gráfica de la función densidad f de una distribución normal (con media $m = np = 2$ y varianza $\sigma^2 = npq = 1{,}6$). Para calcular la probabilidad $P(a \leq t \leq b)$ a partir de la función de masa p sumamos los valores de $p(t)$ en los puntos en que está concentrada la masa del intervalo $a \leq t \leq b$. Cada valor $p(t)$ puede interpretarse como el área de un rectángulo de altura $p(t)$ colocada sobre un intervalo de longitud unidad centrado en t. (Como ejemplo, vemos en la figura 14.11 un rectángulo de ésos centrado en $t = 3$.) La

fórmula (14.25) es el resultado de sustituir las áreas de esos rectángulos por el área del conjunto de ordenadas de f sobre el intervalo $[a - \frac{1}{2}, b + \frac{1}{2}]$.

14.15 Observaciones sobre distribuciones más generales

En las secciones precedentes hemos comentado ejemplos de distribuciones discretas y continuas. Los valores de una distribución discreta se calculan sumando los valores de la función de masa de probabilidad correspondiente. Los de una distribución continua se calculan integrando la función de densidad. Existen, naturalmente, distribuciones que no son discretas ni continuas. Entre éstas están las llamadas de tipo «mixto» en las que la distribución de masas es parcialmente discreta y parcialmente continua. (En la figura 14.3 se ha dibujado un ejemplo.)

Una distribución F se llama *mixta* si puede expresarse como combinación lineal de la forma

$$F(t) = c_1F_1(t) + c_2F_2(t), \tag{14.26}$$

siendo F_1 discreta y F_2 continua. Las constantes c_1 y c_2 deben satisfacer las relaciones

$$0 < c_1 < 1, \qquad 0 < c_2 < 1, \qquad c_1 + c_2 = 1.$$

Las propiedades de las distribuciones mixtas pueden deducirse estudiando las que son discretas o continuas y recurriendo entonces a la linealidad expresada en la igualdad (14.26).

Un tipo de integral, conocida con el nombre de *integral de Riemann-Stieltjes*, hace posible el manejo simultáneo de los casos discreto, continuo y mixto.(*) Si bien esta integral unifica el estudio teórico de las funciones de distribución, en cualquier problema el cálculo de las probabilidades debe reducirse a la integración y a la suma ordinarias. En esta exposición introductoria no pretendemos describir la integral de Riemann-Stieltjes. En consecuencia, gran parte de los temas los discutimos paralelamente, a pares, uno para el caso discreto y otro para el continuo. Sin embargo, daremos únicamente los detalles de manera completa para uno de los casos, dejando para el lector el trabajo de completar los demás.

Aun la misma integral de Riemann-Stieltjes es inadecuada para tratar las funciones de distribución *más generales*. Pero un concepto más potente, la integral de Lebesgue-Stieltjes (*), nos permite tratar satisfactoriamente todos los casos. Sin un conocimiento de este tipo de integral no se puede realizar el estudio de la teoría de probabilidades desde un punto de vista superior.

(*) En el capítulo 9 de la obra del autor *Análisis Matemático*, Editorial Reverté, Barcelona, 1963, puede verse una exposición amplia de la integral de Riemann-Stieltjes.

(*) Véase cualquier libro de teoría de la medida.

14.16 Ejercicios

1. Sea X una variable aleatoria que mide la vida (en horas) de un cierto tipo de tubo de vacío. Supongamos que X tiene una distribución exponencial con parámetro $\lambda = 0{,}001$. Determinar T de modo que $P(X > T)$ sea a) 0,90; b) 0,99. Puede emplearse en el cálculo la fórmula aproximada $-\log(1 - x) = x + x^2/2$.
2. Un material radioactivo sigue una ley de desintegración exponencial con vida media de 2 años. Consideremos el tiempo de desintegración X (en años) de un átomo y supongamos que X es una variable aleatoria con una distribución exponencial. Calcular la probabilidad de que un átomo se desintegre a) en el intervalo $1 \leq X \leq 2$; b) en el intervalo $2 \leq X \leq 3$; c) en el intervalo $2 \leq X \leq 3$, en el supuesto de que no se ha desintegrado en el intervalo $0 \leq X \leq 2$; d) en el intervalo $2 \leq X \leq 3$, en el supuesto de que no se ha desintegrado en el intervalo $1 \leq X \leq 2$.
3. La duración (en minutos) de las llamadas telefónicas a larga distancia desde Caltech se vio que es un fenómeno aleatorio con función densidad de probabilidad

$$f(t) = \begin{cases} ce^{-t/3} & \text{para} \quad t > 0, \\ 0 & \text{para} \quad t \leq 0. \end{cases}$$

Determinar el valor de c y calcular la probabilidad de que una llamada a larga distancia dure a) menos de 3 minutos; b) más de 6 minutos; c) entre 3 y 6 minutos; d) más de 9 minutos.
4. Dadas las constantes reales $\lambda > 0$ y c. Sea

$$f(t) = \begin{cases} \lambda e^{-\lambda(t-c)} & \text{si} \quad t \geq c, \\ 0 & \text{si} \quad t < c. \end{cases}$$

Comprobar que $\int_{-\infty}^{\infty} f(t)\, dt = 1$, y determinar una función de distribución F que tenga f como densidad. Ésta se llama distribución exponencial con dos parámetros, un *parámetro de amortiguamiento* λ y un *parámetro de posición* c.
5. Establecer y demostrar una extensión del teorema 14.9 para distribuciones exponenciales con dos parámetros λ y c.
6. Una variable aleatoria X tiene una distribución exponencial con dos parámetros λ y c. Sea $Y = aX + b$, siendo $a > 0$. Demostrar que Y tiene también una distribución exponencial con dos parámetros λ' y c', y determinar esos parámetros en función de a, b, c y λ.
7. En el ejercicio 16 de la sección 11.28 se demostró que $\int_0^{\infty} e^{-x^2}\, dx = \sqrt{\pi}/2$. Utilizar este resultado para demostrar que para $\sigma > 0$ tenemos

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(\frac{u-m}{\sigma}\right)^2\right\} du = 1 .$$

8. Una variable aleatoria X tiene una distribución normal canónica Φ. Demostrar que a) $\Phi(-x) = 1 - \Phi(x)$; *b)* $P(|X| < k) = 2\Phi(k) - 1$; c) $P(|X| > k) = 2(1 - \Phi(k))$.
9. Una variable aleatoria X tiene una distribución normal canónica Φ. Utilizar la tabla 14.1

para calcular cada una de las siguientes probabilidades: a) $P(X > 0)$; b) $P(1 < X < 2)$; c) $P(X < 3)$; d) $P(|X| > 2)$.

10. Una variable aleatoria X tiene una distribución normal canónica Φ. Utilizar la tabla 14.1 para encontrar un número c tal que a) $P(|X| > c) = \frac{1}{2}$; b) $P(|X| > = 0{,}98$.
11. Supongamos que X tiene una función de distribución normal F con media m y varianza σ^2, y designemos con Φ la distribución normal canónica.

    a) Probar que

$$F(t) = \Phi\left(\frac{t-m}{\sigma}\right).$$

    b) Hallar un valor de c tal que $P(|X - m| > c) = \frac{1}{2}$.
    c) Hallar un valor de c tal que $P(|X - m| > c) = 0{,}98$.
12. Una variable aleatoria está distribuida normalmente con media $m = 1$ y varianza $\sigma^2 = 4$. Calcular cada una de las probabilidades siguientes: a) $P(-3 \leq X \leq 3)$; b) $P(-5 \leq X \leq 3)$.
13. Un arquitecto está diseñando un portal de un edificio que tienen que utilizar personas cuya altura está distribuida normalmente, con media $m = 5$ pies 9 pulgadas, y varianza σ^2 siendo $\sigma = 3$ pulgadas. ¿Qué altura mínima puede dar al portal para que no más del 1 % de las personas choquen con su cabeza en la parte superior del mismo?
14. Si X tiene una distribución normal canónica, demostrar que la variable aleatoria $Y = aX + b$ es también normal si $a \neq 0$. Determinar la media y la varianza de Y.
15. Supongamos que una variable aleatoria X tiene una distribución normal canónica, y pongamos $Y = X^2$.

    a) Demostrar que $F_Y(t) = \dfrac{2}{\sqrt{2\pi}} \displaystyle\int_0^{\sqrt{t}} e^{-u^2/2}\, du \qquad \text{si } t \geq 0.$

    b) Determinar $F_Y(t)$ cuando $t < 0$ y describir la función densidad f_Y.

14.17 Distribuciones de funciones de variables aleatorias

Si φ es una función real cuyo dominio incluye el rango de la variable aleatoria X, podemos construir una nueva variable aleatoria Y mediante la ecuación

$$Y = \varphi(X),$$

lo que significa que $Y(\omega) = \varphi[X(\omega)]$ para cada ω del espacio muestral. Si conocemos la función de distribución F_X de X, ¿cómo encontramos la distribución F_Y de Y? Comenzamos con un caso particular importante. Supongamos que φ es continua y creciente en sentido estricto en todo el eje real. En este caso φ posee una función inversa ψ continua y estrictamente creciente, y es:

$$y = \varphi(x) \qquad \text{si, y sólo si} \qquad x = \psi(y).$$

Según la definición de F_Y tenemos,

$$F_Y(t) = P(Y \leq t) = P[\varphi(X) \leq t].$$

Puesto que φ es estrictamente creciente y continua, los sucesos «$\varphi(x) \leq t$» y «$X \leq \psi(t)$» son idénticos. Por lo tanto, $P[\varphi(X) \leq t] = P[X \leq \psi(t)] = F_X[\psi(t)]$. Luego las distribuciones F_Y y F_X están ligadas por la ecuación

(14.27) $$F_Y(t) = F_X[\psi(t)].$$

Cuando la distribución F_X y la función ψ son derivables podemos derivar ambos miembros de (14.27), utilizando en el segundo miembro la regla de la cadena, para obtener

$$F'_Y(t) = F'_X[\psi(t)] \cdot \psi'(t).$$

Esto nos proporciona la siguiente relación entre las densidades.

$$f_Y(t) = f_X[\psi(t)] \cdot \psi'(t).$$

EJEMPLO 1. $Y = aX + b,\ a > 0$. En este caso tenemos:

$$\varphi(x) = ax + b, \qquad \psi(y) = \frac{y - b}{a}, \qquad \psi'(y) = \frac{1}{a}.$$

Ya que φ es continua y estrictamente creciente podemos escribir:

$$F_Y(t) = F_X\left(\frac{t - b}{a}\right) \qquad \text{y} \qquad f_Y(t) = \frac{1}{a} f_X\left(\frac{t - b}{a}\right).$$

EJEMPLO 2. $Y = X^2$. En este caso $\varphi(x) = x^2$ y el razonamiento anterior no es directamente aplicable pues φ no es estrictamente creciente. No obstante, podemos usar el mismo método para determinar F_Y y f_Y. Por definición tenemos:

$$F_Y(t) = P(X^2 \leq t).$$

Si $t < 0$ el suceso «$X^2 \leq t$» es vacío y por tanto $P(X^2 \leq t) = 0$. Por consiguiente $F_Y(t) = 0$ para $t < 0$. Si $t > 0$ tenemos

$$P(X^2 \leq t) = P(-\sqrt{t} \leq X \leq \sqrt{t}) = F_X(\sqrt{t}) - F_X(-\sqrt{t}) + P(X = -\sqrt{t}).$$

Para una distribución continua F_X tenemos $P(X = -\sqrt{t}) = 0$ y obtenemos la siguiente relación entre F_Y y F_X:

$$F_Y(t) = \begin{cases} 0 & \text{si } t < 0, \\ F_X(\sqrt{t}) - F_X(-\sqrt{t}) & \text{si } t > 0. \end{cases}$$

Para todo $t < 0$ y para aquellos valores de $t > 0$ tales que F_X es derivable en $\sqrt{t}$ y en $-\sqrt{t}$ tenemos la siguiente relación entre las densidades:

$$f_Y(t) = \begin{cases} 0 & \text{si } t < 0, \\ \dfrac{f_X(\sqrt{t}) + f_X(-\sqrt{t})}{2\sqrt{t}} & \text{si } t > 0. \end{cases}$$

Otros problemas de este tipo se discutirán en la sección 14.23 con la ayuda de variables aleatorias bidimensionales.

14.18 Ejercicios

1. Supongamos que X tiene una distribución uniforme en el intervalo $[0, 1]$. Determinar la función de distribución F_Y y la densidad f_Y de la variable aleatoria Y si:

   a) $Y = 3X + 1$,
   b) $Y = -3X + 1$,
   c) $Y = X^2$,
   d) $Y = \log |X|$,
   e) $Y = \log X^2$,
   f) $Y = e^X$.

2. Sea X una variable aleatoria con una función de distribución continua F_X. Si φ es continua y estrictamente creciente en todo eje real y si $\varphi(x) \to a$ cuando $x \to -\infty$ y $\varphi(x) \to b$ cuando $x \to \infty$, determinar la función de distribución F_Y de la variable aleatoria $Y = \varphi(X)$. Calcular también la densidad f_Y, suponiendo que F_X y φ son derivables.
3. Supongamos que X tiene distribución normal canónica. Determinar la función de densidad de la variable aleatoria Y cuando

   a) $Y = X^2$,
   b) $Y = |X|^{1/2}$,
   c) $Y = e^X$,
   d) $Y = \arctan X$.

14.19 Distribución de variables aleatorias bidimensionales

El concepto de distribución puede generalizarse a variables aleatorias n dimensionales. La consideración del caso $n = 2$ nos indicará como puede procederse en los demás.

Si X e Y son dos variables aleatorias unidimensionales definidas en un espacio muestral común S, (X, Y) representará la variable aleatoria bidimensional cuyo valor en un punto genérico ω de S queda determinado por el par de números reales $(X(\omega), Y(\omega))$. La notación

$$X \leq a, \ Y \leq b$$

es una forma abreviada de designar el conjunto de todos los elementos ω de S tales

que $X(\omega) \leq a$ e $Y(\omega) \leq b$; la probabilidad de este suceso se representa por

$$P(X \leq a,\ Y \leq b).$$

Análogamente se definen las notaciones $a < X \leq b$, $c < Y \leq d$, y $P(a < X \leq b,\ c < Y \leq d)$.

El conjunto de puntos (x, y) tales que $x \leq a$ e $y \leq b$ es el producto cartesiano $A \times B$ de los dos intervalos infinitos unidimensionales $A = \{x \mid x \leq a\}$ y $B = \{y \mid y \leq b\}$. El conjunto $A \times B$ se representa geométricamente por la región rectangular infinita dibujada en la figura 14.12. El número $P(X \leq a,\ Y \leq b)$ representa la probabilidad de que un punto $(X(\omega), Y(\omega))$ esté situado en esa región. Estas probabilidades se utilizan para definir las distribuciones de probabilidad bidimensionales.

DEFINICIÓN. *La función de distribución de la variable aleatoria bidimensional (X, Y) es la función real F definida para todo par de números reales a y b mediante la igualdad*

$$F(a, b) = P(X \leq a,\ Y \leq b).$$

que también se conoce como distribución conjunta de las dos variables unidimensionales X e Y.

Para calcular la probabilidad de que (X, Y) esté en un rectángulo, empleamos el siguiente teorema, generalización del teorema 14.2 b).

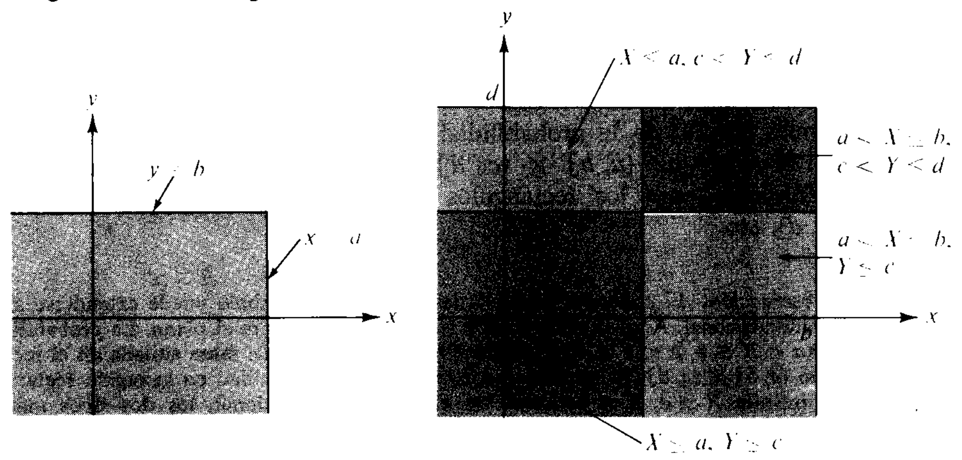


FIGURA 14.12 *Región rectangular infinita $A \times B$, siendo $A = \{x \mid x \leq a\}$ y $B = \{y \mid y \leq b\}$.*

FIGURA 14.13 *El suceso «$X \leq b,\ Y \leq d$» expresado como reunión de cuatro sucesos disjuntos.*

TEOREMA 14.10. *Si F es la función de distribución de una variable aleatoria bidimensional (X, Y) y si $a<b$ y $c<d$, entonces*

$$(14.28) \quad P(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

Demostración. Los dos sucesos «$X \leq a, c < Y \leq d$» y «$X \leq a, Y \leq c$» son disjuntos, y su reunión es «$X \leq a, Y \leq d$». Sumando las probabilidades obtenemos $P(X \leq a, c < Y \leq d) + P(x \leq a, Y \leq c) = P(X \leq a, Y \leq d)$; luego,

$$P(X \leq a, c < Y \leq d) = F(a, d) - F(a, c).$$

Análogamente, tenemos:

$$P(a < X \leq b, Y \leq c) = F(b, c) - F(a, c).$$

Los cuatro sucesos

$$\text{"}X \leq a, Y \leq c,\text{"} \qquad \text{"}X \leq a, c < Y \leq d,\text{"}$$

$$\text{"}a < X \leq b, Y \leq c,\text{"} \qquad \text{"}a < X \leq b, c < Y \leq d\text{"}$$

son disjuntos, y su reunión es «$X \leq b, Y \leq d$». (Véase figura 14.13.) Sumando las correspondientes probabilidades y en virtud de las dos igualdades precedentes obtenemos:

$$F(a, c) + [F(a, d) - F(a, c)] + [F(b, c) - F(a, c)] +$$
$$+ P(a < X \leq b, c < Y \leq d) = F(b,d),$$

que equivale a (14.28).

La fórmula (14.28) da la probabilidad de que la variable (X, Y) tenga un valor interior al rectángulo $(a, b] \times (c, d]$. Existen, como es natural, las correspondientes fórmulas para los rectángulos $[a, b] \times [c, d]$, $(a, b) \times (c, d)$, $[a, b) \times [c, d)$, etc.

Observación. El concepto de masa de probabilidad también puede extenderse al caso bidimensional. Aquí la masa total 1 está distribuida sobre el plano. La probabilidad $P(a < X \leq b, c < Y \leq d)$ representa la cantidad total de masa situada en el rectángulo $(a, b] \times (c, d]$. El número $F(a, b)$ representa la cantidad en la región rectangular infinita $X \leq a$, $Y \leq b$. Como en el caso unidimensional, los dos tipos más importantes de distribuciones son las *discretas* y las *continuas*. En el caso discreto toda la masa está concentrada en un número finito o en una infinidad numerable de puntos. En el caso continuo la masa está repartida sobre el plano con un espesor uniforme o variable.

14.20 Distribuciones discretas bidimensionales

Dada la variable aleatoria (X, Y) podemos definir una nueva función p, llamada la función de masa de probabilidad de (X, Y), tal que

$$p(x, y) = P(X = x,\ Y = y)$$

para todo par de números reales (x, y). Sea T el conjunto de puntos (x, y) para los cuales $p(x, y) > 0$. Puede demostrarse que T es finito o infinito numerable. Si la suma de las probabilidades $p(x, y)$ para todo (x, y) de T es igual a 1, esto es, si

(14.29) $$\sum_{(x,y)\in T} p(x, y) = 1\,,$$

la variable aleatoria (X, Y) se llama *discreta.* Los puntos (x, y) de T son *puntos de masa* de (X, Y).

Supongamos que $x_1, x_2, x_3, \ldots$ e $y_1, y_2, y_3, \ldots$ son valores posibles de X e Y respectivamente, y sea

$$p_{ij} = P(X = x_i\,,\ Y = y_j)\,.$$

Si cada p_{ij} es positiva y la suma de todas las p_{ij} es 1, la probabilidad del suceso «$(X, Y) \in E$» es la suma de todas las p_{ij} considerando todos los x_i e y_j para los que $(x_i, y_j) \in E$. Indicamos esto escribiendo

$$P[(X, Y) \in E] = \sum_{\substack{x_i \quad y_j \\ (x_i,y_j)\in E}}\sum p_{ij}\,.$$

En particular, puesto que $P(X \leq x,\ Y \leq y) = F(x, y)$, la distribución conjunta F (que también se llama discreta) viene dada por la doble suma

(14.30) $$F(x, y) = \sum_{x_i \leq x}\ \sum_{y_j \leq y} p_{ij}\,.$$

Los números p_{ij} también pueden usarse para reconstruir las funciones de masa de probabilidad p_X y p_Y de las variables aleatorias unidimensionales X e Y. En efecto, si E_{ij} representa el suceso «$X = x_i,\ Y = y_j$», los sucesos E_{i1}, E_{i2}, $E_{i3}, \ldots$

son disjuntos y su reunión es el suceso «$X = x_i$». Luego, por aditividad numerable, obtenemos

$$P(X = x_i) = \sum_{j=1}^{\infty} P(E_{ij}) = \sum_{j=1}^{\infty} p_{ij}\,. \tag{14.31}$$

Análogamente, obtenemos

$$P(Y = y_j) = \sum_{i=1}^{\infty} P(E_{ij}) = \sum_{i=1}^{\infty} p_{ij}\,. \tag{14.32}$$

Por lo tanto, las correspondientes distribuciones unidimensionales F_X y F_Y pueden calcularse a partir de las fórmulas

$$F_X(t) = \sum_{x_i \leq t} P(X = x_i) = \sum_{x_i \leq t} \sum_{j=1}^{\infty} p_{ij}$$

y

$$F_Y(t) = \sum_{y_j \leq t} P(Y = y_j) = \sum_{y_j \leq t} \sum_{i=1}^{\infty} p_{ij}\,.$$

Para espacios muestrales finitos, naturalmente, las series deben reemplazarse por sumas finitas.

14.21 Distribuciones continuas bidimensionales. Funciones de densidad

Como podía suponerse, las *distribuciones continuas* son las que son continuas en todo el plano. Para la mayoría de las distribuciones continuas F que se presentan en la práctica existe una función no negativa f (que se llama *densidad de probabilidad* de F) tal que las probabilidades de la mayor parte de los sucesos que nos interesan pueden calcularse por doble integración de la densidad. Esto es, la probabilidad de un suceso «$(X, Y) \in Q$» es:

$$P[(X, Y) \in Q] = \iint_Q f. \tag{14.33}$$

Cuando tal función f existe se llama también densidad de probabilidad de la variable aleatoria (X, Y), o densidad conjunta de X e Y. No intentaremos describir la clase de regiones Q para las que (14.33) es válida, tan sólo mencionaremos que esta clase es lo bastante amplia para incluir todas las regiones que se presentan en las aplicaciones ordinarias de las probabilidades. Por ejemplo, si existe una densidad conjunta tenemos siempre

$$P(a < X \leq b, c < Y \leq d) = \iint_R f(x, y)\, dx\, dy\,, \tag{14.34}$$

donde $R = [a, b] \times [c, d]$. El integrando f frecuentemente es tal que la integral doble se puede calcular por integración unidimensional repetida en cuyo caso (14.34) se transforma en

$$P(a < X \leq b, c < Y \leq d) = \int_c^d \left[\int_a^b f(x, y)\, dx\right] dy = \int_a^b \left[\int_c^d f(x, y)\, dy\right] dx .$$

En todos los ejemplos que consideraremos, esta fórmula es también válida en los casos límite en los que a y c se sustituyen por $-\infty$ y b y d por $+\infty$. Así pues, tenemos

$$(14.35) \qquad F(b, d) = \int_{-\infty}^d \left[\int_{-\infty}^b f(x, y)\, dx\right] dy = \int_{-\infty}^b \left[\int_{-\infty}^d f(x, y)\, dy\right] dx$$

para todo par b y d, y

$$(14.36) \qquad \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(x, y)\, dx\right] dy = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(x, y)\, dy\right] dx = 1 .$$

Las igualdades (14.35) y (14.36) son las correspondientes a las (14.30) y (14.29) de los casos discretos, respectivamente.

Si existe una densidad no es única, ya que el integrando en (14.33) puede modificarse en un número finito de puntos sin afectar el valor de la integral. Sin embargo, existe a lo más una función de densidad continua. En efecto, en los puntos de continuidad de f tenemos las fórmulas

$$f(x, y) = D_{1,2}F(x, y) = D_{2,1}F(x, y) ,$$

obtenidas por derivación de las integrales (14.35).

Como en el caso discreto, la densidad conjunta f puede usarse para calcular las densidades unidimensionales f_X y f_Y. Las fórmulas análogas a las (14.31) y (14.32) son

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy \qquad \text{y} \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)\, dx .$$

Las correspondientes distribuciones $F_X(t)$ y $F_Y(t)$ se obtienen, naturalmente, integrando las respectivas densidades f_X y f_Y entre $-\infty$ y t.

Las variables aleatorias X e Y se llaman *independientes* si la distribución conjunta $F(x, y)$ puede factorizarse como sigue

$$F(x, y) = F_X(x)F_Y(y)$$

para todo (x, y). En la próxima sección de ejercicios se discutirán algunas consecuencias de la independencia.

EJEMPLO. Consideremos la función f que tiene el valor constante 1 sobre el cuadrado $R = [0, 1] \times [0, 1]$, y el valor 0 en el resto del plano. Una variable aleatoria (X, Y) que tenga esta función de densidad se llama *distribuida uniformemente* sobre R. La correspondiente función de distribución F es:

$$F(x,y) = \begin{cases} xy & \text{si} \quad (x,y) \in R, \\ x & \text{si} \quad 0 < x < 1 \quad \text{y} \quad y > 1, \\ y & \text{si} \quad 0 < y < 1 \quad \text{y} \quad x > 1, \\ 1 & \text{si} \quad x \geq 1 \quad \text{y} \quad y \geq 1, \\ 0 & \text{en los demás puntos.} \end{cases}$$

La gráfica de F sobre R es una parte de la superficie $z = xy$ (silla de montar). En los puntos interiores de R existen las derivadas parciales mixtas $D_{1,2}F(x, y)$ y $D_{2,1}F(x, y)$ y son iguales a $f(x, y)$. Esta distribución es el producto de dos distribuciones uniformes unidimensionales F_X y F_Y. Luego X e Y son independientes.

14.22 Ejercicios

1. Sean X e Y dos variables aleatorias unidimensionales con distribuciones F_X y F_Y, y sea F la distribución conjunta de X e Y.
   a) Demostrar que X e Y son independientes si, y sólo si, tenemos

$$P(a < X \leq b, c < Y \leq d) = P(a < X \leq b)P(c < Y \leq d)$$

   cualesquiera que sean a, b, c, d, con $a < b$ y $c < d$.
   b) Consideremos el caso discreto. Supongamos que $x_1, x_2, \ldots$ e $y_1, y_2 \ldots$ son los puntos de masa de X e Y, respectivamente. Sean $a_i = P(X = x_i)$ y $b_j = P(Y = y_j)$. Si $p_{ij} = P(X = x_i,\ Y = y_j)$, demostrar que X e Y son independientes si $p_{ij} = a_i b_j$ para todo par i, j.
   c) Supongamos que X e Y tienen distribuciones continuas con las densidades f_X y f_Y y sea f la densidad de la distribución conjunta. Admitamos la continuidad de las tres densidades. Demostrar que la condición de independencia es equivalente a la afirmación $f(x, y) = f_X(x) f_Y(y)$ para todo (x, y). [*Indicación:* Expresar f como derivada de la distribución conjunta F.]
2. En el ejercicio 1 suponer que $P(X = x_1,\ Y = y_1) = P(X = x_2,\ Y = y_2) = p/2$ y que $P(X = x_1,\ Y = y_2) = P(X = x_2,\ Y = y_1) = q/2$, siendo p y q no negativas de suma 1.
   a) Determinar las probabilidades unidimensionales $P(X = x_i)$ y $P(Y = y_j)$ para $i = 1, 2$ y $j = 1, 2$.
   b) ¿Para qué valor (o valores) de p serán independientes X e Y?

3. Si $a < b$ y $c < d$, definir f del modo siguiente:

$$f(x,y) = \begin{cases} \dfrac{1}{(b-a)(d-c)} & \text{si } (x,y) \in [a,b] \times [c,d], \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

a) Comprobar que ésta es la densidad de una distribución continua F y determinarla.
b) Determinar las distribuciones unidimensionales F_X y F_Y.
c) Determinar si X e Y son o no independientes.

4. Si $P(Y \le b) \ne 0$, la probabilidad condicionada de que $X \le a$, dado que $Y \le b$, se representa con el símbolo $P(X \le a \mid Y \le b)$, y se define mediante la igualdad

$$P(X \le a \mid Y \le b) = \frac{P(X \le a,\ Y \le b)}{P(Y \le b)}.$$

Si $P(Y \le b) = 0$, definimos $P(X \le a \mid Y \le b) = P(X \le a)$. Análogamente, si $P(X \le a) \ne 0$, definimos $P(Y \le b \mid X \le a) = P(X \le a, Y \le b)/P(X \le a)$. Si $P(X \le a) = 0$, definimos $P(Y \le b \mid X \le a) = P(Y \le b)$.

a) Probar la independencia de X e Y y expresarla en función de probabilidades condicionadas.

b) Considerar el caso discreto. Supongamos que $x_1, x_2, \ldots$ e $y_1, y_2, \ldots$ son los puntos de masa de X e Y respectivamente. Demostrar que

$$P(X = x_i) = \sum_{j=1}^{\infty} P(Y = y_j) P(X = x_i \mid Y = y_j)$$

y

$$P(Y = y_j) = \sum_{i=1}^{\infty} P(X = x_i) P(Y = y_j \mid X = x_i).$$

5. Una casa de juego ofrece a sus clientes el juego siguiente: Se lanza una moneda. Si el resultado de la primera tirada es cruz, el jugador pierde y el juego vuelve a empezar. Si en cambio sale cara, vuelve a echarse la moneda. Si en la segunda tirada aparece cara el jugador gana 2 dólares, pero si sale cruz gana sólo 1 dólar. Sea X la variable aleatoria que vale 1 ó 0 según que en la primera tirada aparezca cara o cruz. Sea Y otra variable aleatoria que cuenta el número de dólares ganados por el jugador. Hacer uso del ejercicio 4 (o de algún otro método) para calcular $P(Y = 0)$, $P(Y = 1)$, y $P(Y = 2)$.

6. Haciendo referencia al ejercicio 4, deducir la llamada fórmula de Bayes:

$$P(X = x_k \mid Y = y_j) = \frac{P(X = x_k) P(Y = y_j \mid X = x_k)}{\sum_{i=1}^{\infty} P(X = x_i) P(Y = y_j \mid X = x_i)},$$

$$P(Y = y_k \mid X = x_i) = \frac{P(Y = y_k) P(X = x_i \mid Y = y_k)}{\sum_{j=1}^{\infty} P(Y = y_j) P(X = x_i \mid Y = y_j)}.$$

7. Dadas dos urnas A y B. La urna A contiene un billete de 5 dólares y dos de 10 dólares. La urna B contiene tres billetes de 5 dólares y uno de 10 dólares. Se extrae un

billete de A y se introduce en B. Sea Y la variable aleatoria que cuenta el valor del billete transferido. A continuación se extrae un billete de la urna B y usamos la variable aleatoria X para contar *su* valor. Calcular las probabilidades condicionadas

$$P(Y = 5 \mid X = 10) \qquad \text{y} \qquad P(Y = 10 \mid X = 10).$$

[*Indicación:* Hacer uso de la fórmula de Bayes del ejercicio 6.]

8. Dadas tres cajas idénticas, cada una de las cuales tiene dos departamentos. La caja número 1 tiene una moneda de oro en un departamento y una de plata en el otro. La caja 2 tiene una moneda de oro en cada departamento y la caja 3 una moneda de plata en cada departamento. Se abre al azar un departamento y se encuentra una moneda de oro. Calcular la probabilidad de que el otro departamento de la misma caja contenga una moneda de plata. [*Indicación:* Emplear la fórmula de Bayes.]
9. Sea Q una región plana y $a(Q)$ su área. Se dice que una variable aleatoria bidimensional (X, Y) tiene *distribución uniforme* sobre Q si su función de densidad f viene definida por las fórmulas:

$$f(x, y) = \begin{cases} 1/a(Q) & \text{si } (x, y) \in Q, \\ 0 & \text{si } (x, y) \notin Q. \end{cases}$$

a) Si E es una subregión de Q con área $a(E)$, demostrar que $a(E)/a(Q)$ es la probabilidad del suceso $(X, Y) \in E$.

b) Caen al azar gotas de agua sobre el cuadrado Q de vértices $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$. Un resultado es el punto (x, y) de Q alcanzado por una determinada gota. Sean $X(x, y) = x$ e $Y(x, y) = y$ y supongamos que (X, Y) tiene distribución uniforme sobre Q. Determinar la función de densidad conjunta f y las densidades unidimensionales f_X y f_Y. ¿Son independientes las variables X e Y?

10. Una variable aleatoria bidimensional (X, Y) tiene la función de distribución conjunta F. Sean $U = X - a$, $V = Y - b$, siendo a y b constantes. Si G representa la distribución conjunta de (U, V) demostrar que

$$G(u, v) = F(u + a, v + b).$$

Deducir una relación semejante entre las funciones de densidad f de (X, Y) y g de (U, V) cuando f es continua.

14.23 Distribuciones de funciones de dos variables aleatorias

Consideremos ahora un problema más general. Si X e Y son variables aleatorias unidimensionales con distribuciones conocidas, ¿cómo encontramos la distribución de nuevas variables aleatorias tales como $X + Y$, XY o $X^2 + Y^2$? El resto de esta sección se dedica a exponer un método que nos ayuda a contestar preguntas como ésta. Definimos dos nuevas variables aleatorias U y V con ecuaciones de la forma

$$U = M(X, Y), \qquad V = N(X, Y),$$

donde $M(X, Y)$ o $N(X, Y)$ es la combinación particular en la que estamos interesados. A partir del conocimiento de la distribución conjunta f de la variable aleatoria bidimensional (X, Y) calculamos la distribución conjunta g de (U, V). Una vez se conoce g, las distribuciones de U y de V se encuentran fácilmente.

Para describir el método con detalle, consideremos una aplicación uno a uno del plano xy sobre el plano uv definida por el par de ecuaciones

$$u = M(x, y), \qquad v = N(x, y).$$

La aplicación inversa sea la definida por

$$x = Q(u, v), \qquad y = R(u, v),$$

y supongamos que Q y R tienen derivadas parciales continuas. Si T es una región en el plano xy, sea T' su imagen en el plano uv, como se ve en la figura 14.14. Sean X e Y dos variables aleatorias unidimensionales continuas que tengan una distribución conjunta continua y sea f la función de densidad de probabilidad de (X, Y). Definamos las nuevas variables aleatorias U y V escribiendo $U = M(X, Y)$, $V = N(X, Y)$. Para determinar la densidad g de la variable aleatoria (U, V) procedemos como sigue:

Las variables aleatorias X e Y están asociadas a un espacio muestral S. Para cada ω de S tenemos $U(\omega) = M[X(\omega), Y(\omega)]$, y $V(\omega) = N[(\omega), Y(\omega)]$. Puesto que la aplicación es uno a uno, los dos conjuntos

$$\{\omega \mid (U(\omega), V(\omega)) \in T'\} \qquad \text{y} \qquad \{\omega \mid (X(\omega), Y(\omega)) \in T\}$$

son iguales. Por consiguiente,

$$P[(U, V) \in T'] = P[(X, Y) \in T]. \tag{14.37}$$

Ya que f es la función de densidad (X, Y) podemos escribir

$$P[(X, Y) \in T] = \iint_T f(x, y)\, dx\, dy. \tag{14.38}$$

En virtud de (14.37) y de la fórmula de transformación de las integrales dobles la (14.38) se convierte en:

$$P[(U, Y) \in T'] = \iint_{T'} f[Q(u, v), R(u, v)] \left| \frac{\partial(Q, R)}{\partial(u, v)} \right| du\, dv.$$

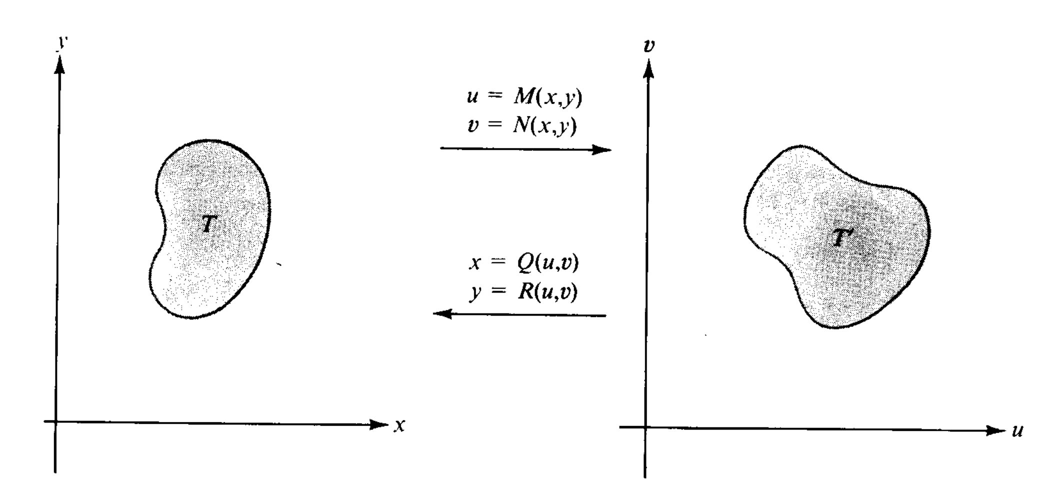


FIGURA 14.14 *Aplicación uno a uno de una región T del plano xy sobre una región T' del plano uv.*

Ya que esto es válido para toda región T' del plano uv la densidad g de (U, V) está definida por el integrando del segundo miembro; esto es, tenemos:

$$g(u, v) = f[Q(u, v), R(u, v)] \left| \frac{\partial(Q, R)}{\partial(u, v)} \right| . \tag{14.39}$$

Las densidades f_U y f_V pueden ahora obtenerse por las fórmulas

$$f_U(u) = \int_{-\infty}^{\infty} g(u, v)\, dv , \qquad f_V(v) = \int_{-\infty}^{\infty} g(u, v)\, du .$$

EJEMPLO 1. *Suma y diferencia de dos variables aleatorias.* Dadas dos variables aleatorias unidimensionales X e Y con densidad conjunta f, determinar las funciones de densidad para las variables aleatorias $U = X + Y$ y $V = X - Y$.

Solución. Utilicemos la aplicación dada por $u = x + y$, $v = x - y$. Esta es una transformación lineal no singular cuya inversa viene dada por

$$x = \frac{u + v}{2} = Q(u, v), \qquad y = \frac{u - v}{2} = R(u, v).$$

El jacobiano es

$$\frac{\partial(Q, R)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial Q}{\partial u} & \dfrac{\partial Q}{\partial v} \\ \dfrac{\partial R}{\partial u} & \dfrac{\partial R}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2} .$$

Aplicando la ecuación (14.39) vemos que una densidad conjunta g de (U, V) viene dada por la fórmula

$$g(u, v) = f\left(\frac{u+v}{2}, \frac{u-v}{2}\right) \left|\frac{\partial(Q, R)}{\partial(u, v)}\right| = \frac{1}{2} f\left(\frac{u+v}{2}, \frac{u-v}{2}\right).$$

Para obtener una densidad $f_U = f_{X+Y}$ integramos respecto a v y encontramos

$$f_{X+Y}(u) = \frac{1}{2}\int_{-\infty}^{\infty} f\left(\frac{u+v}{2}, \frac{u-v}{2}\right) dv.$$

El cambio de variable $x = \frac{1}{2}(u + v)$, $dx = \frac{1}{2}dv$, transforma esa fórmula en la siguiente:

$$f_{X+Y}(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx.$$

Del mismo modo, encontramos

$$f_{X-Y}(v) = \frac{1}{2}\int_{-\infty}^{\infty} f\left(\frac{u+v}{2}, \frac{u-v}{2}\right) du = \int_{-\infty}^{\infty} f(x, x - v)\, dx.$$

Un caso particular importante se presenta cuando X e Y son independientes. En este caso la densidad de probabilidad conjunta se escinde en un producto

$$f(x, y) = f_X(x) f_Y(y),$$

y las integrales que dan f_{X+Y} y f_{X-Y} se convierten en

$$f_{X+Y}(u) = \int_{-\infty}^{\infty} f_X(x) f_Y(u - x)\, dx, \qquad f_{X-Y}(v) = \int_{-\infty}^{\infty} f_X(x) f_Y(x - v)\, dx.$$

EJEMPLO 2. *Suma de dos distribuciones exponenciales.* Supongamos ahora que cada una de las variables X e Y tiene una distribución exponencial, por ejemplo $f_X(t) = f_Y(t) = 0$ para $t < 0$, y

$$f_X(t) = \lambda e^{-\lambda t}, \qquad f_Y(t) = \mu e^{-\mu t} \qquad \text{para } t \geq 0.$$

Determinar la densidad de $X + Y$ cuando X e Y son independientes.

Solución. Si $u < 0$ la integral que da $f_{X+Y}(u)$ es 0 ya que el factor $f_X(x) = 0$

para $x < 0$, y el factor $f_Y(u - x) = 0$ para $x \geq 0$. Si $u \geq 0$ la integral que da $f_{X+Y}(u)$ se transforma en

$$f_{X+Y}(u) = \int_0^u \lambda e^{-\lambda x} \mu e^{-\mu(u-x)}\, dx = \lambda\mu e^{-\mu u} \int_0^u e^{(\mu-\lambda)x}\, dx .$$

Para calcular la última integral consideremos dos casos, $\mu = \lambda$ y $\mu \neq \lambda$.

Si $\mu = \lambda$ la integral tiene el valor u y obtenemos

$$f_{X+Y}(u) = \lambda^2 u e^{-\lambda u} \quad \text{para} \quad u \geq 0 .$$

Si $\mu \neq \lambda$ obtenemos

$$f_{X+Y}(u) = \lambda\mu e^{-\mu u} \frac{e^{(\mu-\lambda)u} - 1}{\mu - \lambda} = \lambda\mu \frac{e^{-\lambda u} - e^{-\mu u}}{\mu - \lambda} \qquad \text{para } u \geq 0 .$$

EJEMPLO 3. *Máximo y mínimo de dos variables aleatorias independientes.* Sean X e Y dos variables aleatorias unidimensionales con densidades f_X y f_Y y funciones de distribución correspondientes F_X y F_Y. Sean U y V las variables aleatorias

$$U = \max\{X, Y\}, \qquad V = \min\{X, Y\} .$$

Esto es, para cada ω en el espacio muestral, $U(\omega)$ es el máximo y $V(\omega)$ el mínimo de los dos números $X(\omega)$, $Y(\omega)$. La aplicación $u = \max\{x, y\}$, $v = \min\{x, y\}$ no es uno a uno, así que el método utilizado para deducir la ecuación (14.39) no es aplicable. Sin embargo, en este caso podemos obtener las funciones de distribución de U y V directamente.

Observemos primero que $U \leq t$ si, y sólo si, $X \leq t$ e $Y \leq t$. Por consiguiente $P(U \leq t) = P(X \leq t,\ Y \leq t)$. En virtud de la independencia esto es igual a $P(X \leq t)P(Y \leq t) = F_X(t)F_Y(t)$. Así pues, tenemos

$$F_U(t) = F_X(t)F_Y(t) .$$

En cada punto de continuidad de f_X y f_Y podemos derivar esa relación obteniendo

$$f_U(t) = f_X(t)F_Y(t) + F_X(t)f_Y(t) .$$

Del mismo modo, tenemos $V > t$ si y sólo si $X > t$ e $Y > t$. Por consiguiente

$$\begin{aligned} F_V(t) &= P(V \leq t) = 1 - P(V > t) = 1 - P(X > t,\ Y > t) = 1 - P(X > t)P(Y > t) \\ &= 1 - (1 - F_X(t))(1 - F_Y(t)) = F_X(t) + F_Y(t) - F_X(t)F_Y(t) . \end{aligned}$$

En los puntos de continuidad de f_X y f_Y derivamos esa relación y obtenemos

$$f_V(t) = f_X(t) + f_Y(t) - f_X(t)F_Y(t) - F_X(t)f_Y(t).$$

14.24 Ejercicios

1. Sean X e Y dos variables aleatorias unidimensionales independientes, con distribución uniforme cada una sobre el intervalo $[0, 1]$. Sean $U = X + Y$ y $V = X - Y$.
   a) Demostrar que U tiene una densidad continua f_U dada por

$$f_U(u) = \begin{cases} u & \text{si } \; 0 < u \leq 1, \\ 2 - u & \text{si } \; 1 < u < 2, \\ 0 & \text{para los demás valores de } u \end{cases}$$

   b) Describir, en forma análoga, una densidad continua f_V para V.
   c) Determinar si U y V son o no independientes.
2. Sean X e Y las del ejercicio 1, y $U = \max\{X, Y\}$, $V = \min\{X, Y\}$.
   a) Demostrar que U tiene una función de densidad tal que $f_U(t) = 2t$ para $0 \leq t < 1$, y $f_U(t) = 0$ para los demás valores de t.
   b) Describir una función de densidad f_V para V.
   c) Determinar si U y V son o no independientes.
3. Sean X e Y dos variables aleatorias unidimensionales independientes, cada una con distribución exponencial con parámetro $\lambda = 1$, y sea $f(x, y) = f_X(x)f_Y(y)$, el producto de las densidades de X e Y.
   a) Sea A el conjunto de puntos del plano xy en los que $f(x, y) > 0$. Dibujar A y su imagen A' en la aplicación definida por $u = x + y$, $v = x/(x + y)$.
   b) Sean $U = X + Y$ y $V = X/(X + Y)$ dos nuevas variables aleatorias. Calcular la densidad g de (U, V).
   c) Calcular la densidad f_U.
   d) Calcular la densidad f_V.
4. Sean X e Y dos variables aleatorias independientes, ambas con distribución normal canónica (media $= 0$, varianza $= 1$). Introduzcamos las nuevas variables aleatorias U y V mediante las ecuaciones $U = X/Y$, $V = Y$. Sea g la función densidad de (U, V).
   a) Demostrar que

$$g(u, v) = -\frac{v}{2\pi} e^{-(1+u^2)v^2/2} \qquad \text{si } \; v < 0.$$

   b) Encontrar una fórmula parecida para calcular $g(u, v)$ cuando $v \geq 0$.
   c) Determinar la función densidad de U.

5. Supongamos que X tenga la función densidad dada por

$$f_X(x) = \begin{cases} \dfrac{1}{\pi\sqrt{1 - x^2}} & \text{si} \quad -1 < x < 1\,, \\ 0 & \text{si} \quad |x| \geq 1\,. \end{cases}$$

Si una variable aleatoria independiente Y tiene la densidad

$$f_Y(y) = \begin{cases} ye^{-y^2/2} & \text{si} \quad y \geq 0\,, \\ 0 & \text{si} \quad y < 0\,, \end{cases}$$

encontrar la densidad de $Z = XY$.

6. Dadas dos variables aleatorias unidimensionales independientes X e Y con densidades continuas f_X y f_Y. Sean U y V dos variables aleatorias tales que $X = U \cos V$, $Y = U \operatorname{sen} V$, siendo $U > 0$ y $-\pi < V \leq \pi$.
a) Demostrar que U tiene una función de densidad tal que $f_U(u) = 0$ para $u < 0$ y

$$f_U(u) = u \int_{-\pi}^{\pi} f_X(u \cos v) f_Y(u \operatorname{sen} v)\, dv \qquad \text{para} \quad u \geq 0\,.$$

b) Determinar f_U y la correspondiente distribución F_U explícitamente cuando cada una de las variables X e Y tiene una distribución normal con media $m = 0$ y varianza σ^2.
7. a) Suponiendo $\sigma_1 > 0$ y $\sigma_2 > 0$, comprobar la identidad algebraica

$$\left(\frac{x - m_1}{\sigma_1}\right)^2 + \left(\frac{t - x - m_2}{\sigma_2}\right)^2 = \left(\frac{x - m_0}{\sigma_0}\right)^2 + \left(\frac{t - (m_1 + m_2)}{\sigma}\right)^2,$$

en donde

$$\sigma^2 = \sigma_1^2 + \sigma_2^2, \qquad \sigma_0^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma^2}, \qquad \text{y} \qquad m_0 = \frac{m_1\sigma_2^2 + (t - m_2)\sigma_1^2}{\sigma_1^2 + \sigma_2^2}.$$

b) Dadas dos variables aleatorias unidimensionales independientes X e Y. Supóngase que X tiene una distribución normal con media m_1 y varianza σ_1^2, y que Y tiene una distribución normal con media m_2 y varianza σ_2^2. Demostrar que $X + Y$ tiene una distribución normal con media $m = m_1 + m_2$ y varianza $\sigma^2 = \sigma_1^2 + \sigma_2^2$.
8. Dadas dos variables aleatorias unidimensionales X e Y con densidades f_X y f_Y y densidad conjunta f. Para cada y fijo, se define

$$f_X(x \mid Y = y) = \frac{f(x, y)}{f_Y(y)} \quad \text{siempre que} \quad f_Y(y) > 0\,.$$

Esta es la llamada densidad de probabilidad condicional de X, supuesto que $Y = y$.

Análogamente, definimos la densidad de probabilidad condicional de Y, en el supuesto de que $X = x$, mediante la ecuación

$$f_Y(y \mid X = x) = \frac{f(x,y)}{f_X(x)} \quad \text{siempre que } f_X(x) > 0.$$

a) Si f_Y y f_X son positivas, demostrar que $\int_{-\infty}^{\infty} f_X(x \mid Y = y)\, dx = \int_{-\infty}^{\infty} f_Y(y \mid X = x)\, dy = 1$.
b) Si f_Y y f_X son positivas, demostrar que

$$f_X(x) = \int_{-\infty}^{\infty} f_Y(y) f_X(x \mid Y = y)\, dy \qquad \text{y} \qquad f_Y(y) = \int_{-\infty}^{\infty} f_X(x) f_Y(y \mid X = x)\, dx.$$

9. Una variable aleatoria (X, Y) tiene *distribución normal bivariada* si su densidad es

$$f(x, y) = \frac{\sqrt{D}}{2\pi} e^{-Q(x,y)/2},$$

siendo

$$Q(x, y) = A_{11}(x - x_0)^2 + 2A_{12}(x - x_0)(y - y_0) + A_{22}(y - y_0)^2.$$

Q se llama «forma cuadrática». Los números A_{11}, A_{12}, A_{22} son constantes y $A_{11} > 0$. El número $D = A_{11}A_{22} - A_{12}^2$ se llama *discriminante* de Q y se supone positivo. Los números x_0 e y_0 son cualesquiera.
a) Demostrar que $Q(x, y)$ puede reducirse a la suma de cuadrados siguiente:

$$Q(x, y) = A_{11}\left(u + \frac{A_{12}}{A_{11}} v\right)^2 + \frac{D}{A_{11}} v^2, \qquad \text{siendo } u = x - x_0,\, v = y - y_0.$$

b) Definir la integral doble *impropia* $\iint_{-\infty}^{+\infty} f(x, y)\, dx\, dy$ como el límite:

$$\iint_{-\infty}^{+\infty} f(x, y)\, dx\, dy = \lim_{t \to +\infty} \iint_{R(t)} f(x, y)\, dx\, dy,$$

en donde $R(t)$ es el cuadrado $[-t, t] \times [-t, t]$. Demostrar que

$$\iint_{-\infty}^{+\infty} f(x, y)\, dx\, dy = 1.$$

[*Indicación:* Utilizar el apartado a) para transformar la integral doble sobre $R(t)$ en otra sobre el plano uv. Efectuar entonces una transformación lineal de variables para simplificar la integral y, por último, hacer tender t a $+\infty$.]

10. Si una variable aleatoria bidimensional (X, Y) tiene distribución normal bivariada como se ha dicho en el ejercicio 9, demostrar que X e Y son asimismo variables aleatorias unidimensionales normales con medias x_0 e y_0 respectivamente, y con varianzas $\sigma^2(X) = A_{22}/D$, $\sigma^2(Y) = A_{11}/D$.
11. Si (X, Y) tiene distribución normal bivariada, demostrar que la variable aleatoria $Z = X + Y$ tiene distribución normal unidimensional con media $x_0 + y_0$ y varianza $(A_{11} - 2A_{12} + A_{22})/D$.

14.25 Esperanza y varianza

La interpretación como masa de las distribuciones de probabilidad puede llevarse a un estadio más avanzado introduciendo los conceptos de *esperanza* y de *varianza.* Juegan éstos, en Cálculo de probabilidades, el mismo papel que el «centro de gravedad» y el «momento de inercia» en Mecánica. Sin utilizar la integral de Stieltjes hay que dar separadamente las definiciones para los casos discreto y continuo.

DEFINICIONES DE ESPERANZA Y VARIANZA. *Sea X una variable aleatoria unidimensional. La esperanza de X y la varianza de X son números reales que se designan con* $E(X)$ *y* $\mathrm{Var}(X)$ *respectivamente, y se definen como sigue:*

a) *Para una variable aleatoria continua con densidad f_X,*

$$E(X) = \int_{-\infty}^{+\infty} t\, f_X(t)\, dt\,,$$

$$\mathrm{Var}\,(X) = \int_{-\infty}^{+\infty} [t - E(X)]^2 f_X(t)\, dt\,.$$

b) *Para una variable aleatoria discreta con puntos de masa $x_1, x_2, \ldots$ y probabilidades $p_k = P(X = x_k)$, definimos:*

$$E(X) = \sum_{k=1}^{\infty} x_k p_k\,,$$

$$\mathrm{Var}\,(X) = \sum_{k=1}^{\infty} [x_k - E(X)]^2 p_k\,.$$

Observación. $E(X)$ y $\mathrm{Var}(X)$ existen únicamente cuando la integral o la serie en cuestión es *absolutamente* convergente. Se sobreentiende que la serie es una suma finita cuando el espacio muestral es finito; en tal caso $E(X)$ y $\mathrm{Var}(X)$ existen siempre. También existen cuando f_X es 0 en el exterior de algún intervalo finito.

La esperanza matemática $E(X)$ es un valor calculado teóricamente asociado a la variable aleatoria X. En cierta forma, la distribución actúa como si toda la masa estuviera concentrada en un solo punto, $E(X)$. La significación precisa de la

esperanza matemática en la teoría de las probabilidades se discutirá en la sección 14.29 en conexión con la llamada «ley de los grandes números».

En Mecánica, el solo conocimiento del centro de gravedad no nos indica la forma en que la masa está extendida o dispersa en torno a su centro. El «segundo momento» o «momento de inercia» nos da una medida de esa dispersión. En probabilidades, este segundo momento es la varianza. En la sección 14.28 veremos que una varianza pequeña indica que son improbables las desviaciones grandes respecto a los valores esperados.

Si bien la esperanza matemática $E(X)$ puede ser positiva o negativa, la varianza $\text{Var}(X)$ siempre es no negativa. También se utiliza el símbolo σ^2 para representar la varianza. Su raíz cuadrada positiva se llama la *desviación típica* y se representa por σ. La desviación típica es un promedio ponderado; en realidad, σ es la media cuadrática ponderada de la distancia de cada valor de X al valor de $E(X)$. En Mecánica, el concepto análogo es el «radio de giro».

EJEMPLO 1. *Distribución uniforme.* Supongamos que X tenga una distribución uniforme sobre un intervalo $[a, b]$. Entonces $f(t) = 1/(b-a)$ si $a < t < b$, y $f(t) = 0$ para los demás valores de t. Por consiguiente, la esperanza de X es:

$$E(X) = \int_{-\infty}^{+\infty} tf(t)\,dt = \frac{1}{b-a}\int_a^b t\,dt = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2},$$

que es el punto medio del intervalo. Si ponemos m en lugar de $(a+b)/2$ y observamos que $m - a = b - m = (b-a)/2$ encontramos:

$$\text{Var}\,(X) = \frac{1}{b-a}\int_a^b (t-m)^2\,dt = \frac{1}{b-a}\int_{a-m}^{b-m} u^2\,du = \frac{(b-a)^2}{12}.$$

Observemos que la varianza tan sólo depende de la *longitud* del intervalo.

EJEMPLO 2. *Distribución binomial.* Si X tiene distribución binomial con parámetros n y p tenemos:

$$E(X) = \sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k},$$

siendo $q = 1 - p$. Para calcular esta suma, sea

$$f(x, y) = (x+y)^n = \sum_{k=0}^{n}\binom{n}{k} x^k y^{n-k}$$

y teniendo en cuenta que

$$\sum_{k=0}^{n} k\binom{n}{k}x^{k-1}y^{n-k} = \frac{\partial f(x, y)}{\partial x} = n(x + y)^{n-1}.$$

Si multiplicamos ambos miembros de esta última ecuación por x y ponemos $x = p$ e $y = q$, obtenemos $E(X) = np$.

Mediante un argumento similar podemos deducir la fórmula

$$\operatorname{Var}(X) = \sum_{k=0}^{n}(k - np)^2\binom{n}{k}p^k q^{n-k} = npq.$$

En el ejercicio 6 de la sección 14.27 se pide la demostración de esta fórmula.

EJEMPLO 3. *Distribución normal.* Las palabras «media» y «varianza» ya fueron introducidas al describir la distribución normal en la sección 14.14. Tales denominaciones se justifican mediante las fórmulas

$$E(X) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} te^{-[(t-m)/\sigma]^2/2}\,dt = m$$

y

$$\operatorname{Var}(X) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} (t - m)^2 e^{-[(t-m)/\sigma]^2/2}\,dt = \sigma^2.$$

En el ejercicio 7 de la sección 14.27 se piden las demostraciones de esas fórmulas.

Los jugadores emplean con frecuencia el concepto de esperanza para decidir si un determinado juego de azar es favorable o desfavorable. Como ejemplo vamos a considerar el juego de apostar sobre el «rojo» o el «negro» en la ruleta.

EJEMPLO 4. *Ruleta.* Una ruleta recorre los números del 0 al 36. El 0 aparece sobre el fondo gris, la mitad de los 36 números restantes sobre fondo rojo y la otra mitad sobre fondo negro. Los métodos corrientes de apuestas son:

1) Apostar 1 $ sobre un color (rojo o negro). Ganancia posible: 2 $.
2) Apostar 1 $ a un solo número (excluido el 0). Ganancia posible: 36 $.
3) Apostar 1 $ a una docena cualquiera de números (0 excluido). Ganancia posible: 3 $.

Si sale el 0 la casa gana y todos los demás jugadores pierden.

Sea X la variable aleatoria que mide la ganancia que se consigue apostando con el método 1). Los posibles valores de X son $x_1 = -1$ y $x_2 = +1$. Las pro-

babilidades son $P(X = x_1) = 19/37$, $P(X = x_2) = 18/37$. Por lo tanto, la esperanza es

$$E(X) = (-1)\tfrac{19}{37} + (+1)\tfrac{18}{37} = -\tfrac{1}{37}\,;$$

este resultado se interpreta corrientemente como que el juego es desfavorable para todo el que participa en él. La justificación matemática de esta interpretación se consigue con una de las *leyes de los grandes números,* que se discuten en la sección 14.29. El lector puede comprobar que la esperanza tiene el mismo valor usando los métodos de apuesta 2) y 3).

EJEMPLO 6. *Un juego de cara y cruz.* En un juego de cara y cruz hay una probabilidad p de que salga cara (H) y una probabilidad q de que salga cruz (T), siendo $0 \leq p \leq 1$ y $q = 1 - p$. Se lanza la moneda sucesivamente hasta que vuelva a aparecer por segunda vez el primer resultado que se obtuvo; entonces termina el juego. Si el primer resultado es H el jugador gana 1 \$ por cada T que salga hasta que volvamos a obtener H. Por ejemplo, si $HTTTH$ gana 3 \$, en cambio, si HH gana 0 \$. Si el primer resultado es T se aplica la misma regla de juego cambiando H por T. El problema consiste en determinar cuánto debería pagar un jugador para participar en este juego. A tal fin consideramos la variable aleatoria que cuenta el número de dólares ganados por el jugador y calcularemos su esperanza.

Como espacio muestral tomamos la colección de todas las posibles partidas que pueden jugarse. Tal conjunto puede expresarse como la reunión de dos conjuntos A y B, donde

$$A = \{TT, THT, THHT, THHHT, \ldots\} \quad \text{y} \quad B = \{HH, HTH, HTTH, HTTTH, \ldots\}.$$

Los elementos del conjunto A (en el orden citado) los designamos por $a_0, a_1, a_2, \ldots$ y los del conjunto B por $b_0, b_1, b_2, \ldots$ Seguidamente asignamos las probabilidades, como sigue:

$$P(a_n) = p^n q^2 \qquad \text{y} \qquad P(b_n) = q^n p^2 .$$

(Cuando $p = 0$, ponemos $P(a_0) = 1$ y $P(x) = 0$ para cualquier otro x de $A \cup B$. Cuando $q = 0$, ponemos $P(b_0) = 1$ y $P(x) = 0$ para todos los demás x.) En la sección 13.21 se demostró que esta asignación de probabilidades es aceptable.

La variable aleatoria X que nos interesa se define sobre el espacio muestral $A \cup B$ como sigue:

$$X(a_n) = X(b_n) = n \qquad \text{para } n = 0, 1, 2, \ldots .$$

El suceso «$X = n$» consta de las dos partidas a_n y b_n, así que tenemos,

$$P(X = n) = p^n q^2 + q^n p^2,$$

donde p^0 y q^0 se consideran 1 cuando $p = 0$ o $q = 0$. La esperanza de X viene dada por la suma

$$(14.40) \qquad E(X) = \sum_{n=0}^{\infty} nP(X = n) = q^2 \sum_{n=0}^{\infty} np^n + p^2 \sum_{n=0}^{\infty} nq^n .$$

Si $p = 0$ o $q = 0$, obtenemos $E(X) = 0$. En los demás casos podemos calcular las sumas de las series (14.40) observando que para $0 < x < 1$ tenemos,

$$\sum_{n=0}^{\infty} nx^n = x \frac{d}{dx} \sum_{n=0}^{\infty} x^n = x \frac{d}{dx}\left(\frac{1}{1 - x}\right) = \frac{x}{(1 - x)^2} .$$

Aplicando este resultado en (14.40) haciendo $x = p$ y $x = q$ obtenemos, para $0 < p < 1$,

$$E(X) = \frac{q^2 p}{(1 - p)^2} + \frac{p^2 q}{(1 - q)^2} = p + q = 1 .$$

Interpretamos este resultado diciendo que el juego es desfavorable para los que pagan más de 1 $ para participar en él.

Este ejemplo tiene especial interés porque la esperanza $E(X)$ es independiente de p cuando $0 < p < 1$. En otros términos, «cargando» la moneda a favor de cara o de cruz no se altera el valor esperado salvo en el caso extremo en el que se haya «cargado» la moneda de tal modo que nunca salga cara o que nunca salga cruz. Nótese que, como función de p, la esperanza $E(X)$ es discontinua en los puntos $p = 0$ y $p = 1$. En los demás tiene el valor constante 1. Este interesante ejemplo fue sugerido al autor por H. S. Zuckerman.

14.26 Esperanza de una función de una variable aleatoria

Si se relaciona una nueva variable aleatoria Y a la X mediante una ecuación de la forma $Y = \varphi(X)$, su esperanza (en el caso continuo) viene dada por

$$(14.41) \qquad E(Y) = \int_{-\infty}^{+\infty} t f_Y(t)\, dt .$$

La esperanza $E(Y)$ puede calcularse directamente en función de la densidad f_X sin

tener que determinar la densidad de Y. Efectivamente, la siguiente fórmula es equivalente a (14.41):

(14.42) $$E(Y) = \int_{-\infty}^{+\infty} \varphi(t) f_X(t)\, dt.$$

La demostración de (14.42) en el caso más general es difícil y no la intentaremos aquí. No obstante, la demostración para muchos casos particulares interesantes es sencilla. En éstos, φ es derivable y estrictamente creciente sobre todo el eje real. Para una variable aleatoria de distribución continua X con densidad f_X tenemos la fórmula siguiente para la densidad f_Y (deducida en la sección 14.17):

$$f_Y(t) = f_X[\psi(t)] \cdot \psi'(t),$$

siendo ψ la inversa de φ. Si aplicamos esto en (14.41) y efectuamos el cambio de variable $u = \psi(t)$ [de modo que $t = \varphi(u)$], obtenemos

$$E(Y) = \int_{-\infty}^{+\infty} t f_Y(t)\, dt = \int_{-\infty}^{+\infty} t f_X[\psi(t)] \cdot \psi'(t)\, dt = \int_{-\infty}^{+\infty} \varphi(u) f_X(u)\, du,$$

que coincide con (14.42).

Cuando la igualdad (14.42) se aplica a $Y = (X - m)^2$, siendo $m = E(X)$, obtenemos

$$E(Y) = \int_{-\infty}^{+\infty} (t - m)^2 f_X(t)\, dt = \mathrm{Var}\,(X).$$

Esto demuestra que la varianza es a su vez una esperanza. También en el caso discreto es válida una fórmula análoga a la (14.42). Con mayor generalidad, puede demostrarse que

$$E[\varphi(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) f(x, y)\, dx\, dy$$

si (X, Y) es una variable aleatoria continua con densidad conjunta f.

Observación. Para las variables aleatorias bidimensionales, la esperanza y la varianza pueden definirse en forma análoga a la empleada en el caso unidimensional, salvo que se emplean integrales y sumas dobles.

14.27 Ejercicios

1. Se lanza un dado. Representemos con X el número de puntos conseguido. Calcular $E(X)$ y $\mathrm{Var}\,(X)$.
2. Supongamos que X es una variable aleatoria continua con una cierta función de densi-

dad. Sea $Y = (X - m)/\sigma$, donde $m = E(X)$ y $\sigma = \sqrt{\text{Var}(X)}$. Demostrar que $E(Y) = 0$ y $E(Y^2) = 1$.

3. Deducir las siguientes propiedades generales de la esperanza y de la varianza para los casos tanto discretos como continuos.

   a) $E(cX) = cE(X)$, donde c es una constante

   b) $\text{Var}(cX) = c^2 \text{Var}(X)$, donde c es una constante

   c) $E(X + Y) = E(X) + E(Y)$.

   d) $\text{Var}(X) = E(X^2) - [E(X)]^2$.

   e) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2E[(X - E(X))(Y - E(Y))]$.

   f) $E[\varphi_1(X) + \varphi_2(Y)] = E[\varphi_1(X)] + E[\varphi_2(Y)]$. [El apartado c) es un caso particular.]

4. Si X e Y son variables aleatorias independientes, demostrar que

   a) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

   b) $E[\varphi(X) \cdot \psi(Y)] = E[\varphi(X)] \cdot E[\psi(Y)]$.

   c) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes con $E(X_k) = m_k$, demostrar que

$$\text{Var}\left[\sum_{k=1}^{n} (X_k - m_k)\right] = \sum_{k=1}^{n} \text{Var}(X_k - m_k) = \sum_{k=1}^{n} \text{Var}(X_k).$$

5. Sean $X_1, X_2, \ldots, X_n$ n variables aleatorias independientes, con esperanzas iguales, $E(X_k) = m$, y la misma varianza, $\text{Var}(X_k) = \sigma^2$. Sea $\overline{X}$ la media aritmética, $\overline{X} = (1/n)\sum_{i=1}^{n} X_i$. Aplicar los ejercicios 3 y 4 para demostrar que $E(\overline{X}) = m$ y $\text{Var}(\overline{X}) = \sigma^2/n$.

6. a) Si $q = 1 - p$, demostrar la fórmula,

$$\sum_{k=0}^{n} (k - np)^2 \binom{n}{k} p^k q^{n-k} = npq,$$

probando así que $\text{Var}(X) = npq$ para una variable aleatoria X que tenga distribución binomial con parámetros n y p. [*Indicación:* $k^2 = k(k - 1) + k$.]

   b) Si X tiene distribución binomial con parámetros n y p, demostrar que puede expresarse como suma de n variables aleatorias independientes $X_1, X_2, \ldots, X_n$, tomando cada una los valores posibles 0 y 1 con probabilidades p y q respectivamente, y con distribución binomial. Utilizar este resultado y el ejercicio 5 para demostrar que $E(X) = np$ y $\text{Var}(X) = npq$.

7. Determinar la esperanza y la varianza (cuando existan) para una variable aleatoria X que tenga

   a) distribución de Poisson con parámetro λ.

   b) distribución de Cauchy.

   c) distribución exponencial con parámetro λ.

   d) distribución normal.

8. Una variable aleatoria X tiene una función densidad dada por

$$f(t) = \frac{C(r)}{|t|^r} \quad \text{si } |t| > 1, \qquad f(t) = 0 \quad \text{si } |t| \le 1,$$

siendo $r > 1$ y $C(r)$ es independiente de t.
a) Expresar $C(r)$ en función de r y hacer un diseño para indicar la naturaleza de la gráfica de f.
b) Determinar la correspondiente función de distribución F_X e indicar la naturaleza de su gráfica.
c) Calcular $P(X < 5)$ y $P(5 < X < 10)$ en función de r.
d) ¿Para qué valores de r tiene X una esperanza finita? Calcular $E(X)$ en función de r cuando sea finita.
e) ¿Para qué valores de r tiene X varianza finita? Calcularla en función de r cuando sea finita.

9. Un jugador de ruleta lo hace de acuerdo con el siguiente «sistema». Juega en series de tres tiradas. En la primera y segunda tiradas apuesta siempre 1 \$ en rojo. Para la tercera tirada procede así:
a) Si ganó en la primera y en la segunda, no apuesta.
b) Si ganó en la primera o en la segunda perdiendo en la otra, apuesta 1 \$ en color opuesto al que ha salido en la segunda tirada.
c) Si perdió en las dos primeras, apuesta 3 \$ al rojo.
Sean X, Y y Z los resultados (en dólares) de la primera, segunda y tercera tiradas respectivamente. Calcular $E(X)$, $E(Y)$, $E(Z)$ y $E(X + Y + Z)$.

10. (*Problema de Petersburgo*). Un jugador lanza una moneda y gana 1 \$ si a la primera tirada sale cara. Si otra vez sale cara gana otro dólar. Si vuelve a salir cara en la tercera tirada gana 2 \$ (en total ha ganado 4 \$). Si obtiene una sucesión de n caras consecutivas su ganancia acumulada es de 2^{n-1} dólares. El juego termina cuando sale cruz. Sea X el número de dólares ganados en una determinada tirada. Calcular $E(X)$. Visto el resultado, ¿cuánto ha de pagar un jugador a una casa de juego para poder tomar parte en un juego con las reglas anteriores?

11. a) Supongamos que X es una variable aleatoria continua con densidad f_X. Sea $Y = (X - m)/\sigma$, donde $m = E(X)$ y $\sigma = \sqrt{\operatorname{Var}(X)}$. Demostrar que

$$E(e^Y) = e^{-m/\sigma} \int_{-\infty}^{+\infty} e^{t/\sigma} f_X(t)\, dt .$$

b) Sea X una variable aleatoria discreta con distribución de Poisson de parámetro λ. Definir Y como en la parte *a*) y demostrar que

$$E(e^Y) = e^{-\lambda G(\lambda)}, \qquad \text{donde} \quad G(\lambda) = 1 + \frac{1}{\sqrt{\lambda}} - e^{1/\sqrt{\lambda}} .$$

12. Una variable aleatoria X tiene distribución normal canónica. Calcular: a) $E(|X|)$, b) $E(e^X)$, c) $\operatorname{Var}(e^X)$, d) $E(\sqrt{X^2 + Y^2})$. En la parte d), Y también tiene distribución normal canónica, pero es independiente de X.

14.28 Desigualdad de Chebyshev

Como ya se mencionó, un valor pequeño de la varianza significa que es poco probable que una variable aleatoria X se desvíe mucho del valor esperado. Para

precisar más esta afirmación introducimos el valor absoluto $|X - E(X)|$ que mide la distancia entre X y $E(X)$. ¿Qué probabilidad hay de que esa distancia supere a un determinado número? Para contestar esta pregunta tenemos que determinar la probabilidad

$$P[|X - E(X)| > c],$$

donde c es un número positivo dado. En el caso continuo tenemos

$$P[|X - E(X)| > c] = 1 - P[|X - E(X)| \leq c] = 1 - P[E(X) - c \leq X \leq E(X) + c]$$

$$= \int_{-\infty}^{+\infty} f_X(t)\,dt - \int_{E(X)-c}^{E(X)+c} f_X(t)\,dt$$

$$(14.43) \qquad = \int_{-\infty}^{E(X)-c} f_X(t)\,dt + \int_{E(X)+c}^{+\infty} f_X(t)\,dt\,;$$

por consiguiente, el cálculo de esa probabilidad puede llevarse a efecto tan pronto se conozca la densidad f_X. Naturalmente, si f_X es *incógnita* este método no nos da información alguna. No obstante, si la *varianza* se conoce, podemos obtener una cota superior en (14.43). Tal cota superior nos la proporciona el siguiente teorema de P. L. Chebyshev (1821-1894), famoso matemático ruso que aportó notables contribuciones al Cálculo de probabilidades y a otras ramas de la Matemática, en especial a la Teoría de números.

TEOREMA 14.11. DESIGUALDAD DE CHEBYSHEV. *Si X es una variable aleatoria unidimensional con esperanza finita $E(X)$ y con varianza* $\operatorname{Var}(X)$, *entonces para todo número positivo c tenemos*

$$(14.44) \qquad P[|X - E(X)| > c] \leq \frac{\operatorname{Var}(X)}{c^2}\,.$$

Demostración. En el caso continuo tenemos:

$$\operatorname{Var}(X) = \int_{-\infty}^{+\infty} [t - E(X)]^2 f_X(t)\,dt$$

$$\geq \int_{-\infty}^{E(X)-c} [t - E(X)]^2 f_X(t)\,dt + \int_{E(X)+c}^{+\infty} [t - E(X)]^2 f_X(t)\,dt$$

$$\geq c^2\left(\int_{-\infty}^{E(X)-c} f_X(t)\,dt + \int_{E(X)+c}^{+\infty} f_X(t)\,dt\right).$$

En virtud de (14.43), el coeficiente de c^2 en el segundo miembro es $P[|X-E(X)|>c]$. Por consiguiente, dividiendo por c^2 obtenemos (14.44). El caso discreto puede tratarse en forma semejante.

La desigualdad de Chebyshev nos dice que cuanto mayor sea c tanto menor será la probabilidad de que $|X - E(X)| > c$. Dicho de otro modo, no es probable que X se desvíe mucho del valor $E(X)$; lo mismo ocurre si la varianza $\text{Var}(X)$ es pequeña.

Si reemplazamos c por $k\sigma$, siendo $k > 0$ y σ representa la desviación típica $[\sigma = \sqrt{\text{Var}(X)}]$, la desiguadad de Chebyshev se convierte en

$$P[|X - E(X)| > k\sigma] \leq \frac{1}{k^2}.$$

Esto es, la probabilidad de que X difiera de su valor esperado más de k veces la desviación típica no excede de $1/k^2$. Por ejemplo, cuando $k = 10$ esa desigualdad nos dice que la probabilidad $P[|X - E(X)| > 10\sigma]$ no excede a 0,010. Es decir, la probabilidad de que un valor observado de X difiera del valor esperado en más de 10 veces la desviación típica, no supera a 0,010. Análogamente, cuando $k = 3$ encontramos que la probabilidad de que un valor observado difiera de la media en más de tres veces la desviación típica, no supera a 0,111.

La desigualdad de Chebyshev es un teorema general que se aplica a todas las distribuciones. En muchas aplicaciones, cuando se tiene más información acerca de la distribución que se considera, se puede mejorar esa desigualdad. Por ejemplo, si X tiene distribución binomial con parámetros n y p se puede demostrar (utilizando la aproximación normal de la distribución binomial) que para valores grandes de n la probabilidad de que un valor observado difiera de la medida en más de tres desviaciones típicas es próxima a 0,003. (Para este resultado, basta $n \geq 12$.) Esta probabilidad es mucho más pequeña que la 0,111 conseguida con la desigualdad de Chebyshev.

EJEMPLO. *Comprobación de una moneda.* Deseamos comprobar si una moneda es «correcta» o no, lanzándola 10 000 veces y registrando el número de caras. Para una moneda «correcta» la variable aleatoria X que cuenta el número de caras tiene distribución binomial con parámetros $n = 10\,000$ y $p = \frac{1}{2}$. La media de X es $np = 5\,000$ y la desviación típica es $\sigma = \sqrt{npq} = 50$. (Véase el ejemplo 2 en la sección 14.25). Como antes se dijo, la probabilidad para que una variable aleatoria con distribución binomial difiera del valor esperado en más de 3σ es aproximadamente 0,003. Por lo tanto, convendremos en decir que una moneda «no es correcta» si el número de caras en 10 000 tiradas difiere de la media en más de 3σ. Puesto que $E(X) = 5\,000$ y $3\sigma = 150$, podríamos afirmar que la moneda «no es correcta» si el número de caras en 10 000 tiradas es menor que 4 850 o mayor que 5 150.

14.29 Leyes de los grandes números

Al hablar de problemas de monedas, se dice a menudo que la probabilidad de

que salga cara con una moneda perfectamente equilibrada es ½. Esto no significa que si una moneda se lance dos veces salga necesariamente cara una sola vez. Ni que en 1 000 tiradas salgan exactamente 500 caras. Representemos con $h(n)$ el número de caras que se presentan en n tiradas. La experiencia nos demuestra que incluso para n muy grande, la razón $h(n)/n$ no es necesariamente ½. No obstante, la misma experiencia nos dice que esa razón parece *aproximarse* a ½ cuando *n crece,* si bien puede oscilar considerablemente en torno a ½ durante el proceso. Esto nos sugiere la posibilidad de demostrar que

(14.45)
$$\lim_{n\to\infty} \frac{h(n)}{n} = \frac{1}{2}.$$

Desgraciadamente, esto no es posible. Una de las dificultades es que el número $h(n)$ depende no tan sólo de n, sino también de la ejecución del experimento. No tenemos medio de conocer de antemano cómo varía $h(n)$ de un experimento a otro. Pero la dificultad real es que *es* posible (aunque no muy probable) que en alguna determinada experiencia la razón $h(n)/n$ no tienda a ½ en absoluto. Por ejemplo, no hay razón para excluir la posibilidad de que salga cara en *todas* las tiradas de la moneda, en cuyo caso $h(n) = n$ y $h(n)/n \to 1$. Por consiguiente, en lugar de intentar la demostración de la fórmula (14.45), encontraremos más razonable (y más útil) averiguar la probabilidad de que $h(n)/n$ difiera de ½ en una cierta cantidad. Dicho de otro modo, dado un cierto número positivo c, determinar la probabilidad

$$P\left(\left|\frac{h(n)}{n} - \frac{1}{2}\right| > c\right).$$

Introduciendo una conveniente variable aleatoria y aplicando la desigualdad de Chebyshev podemos conseguir una *cota superior* útil de esa probabilidad, una cota que no exige un conocimiento explícito de $h(n)$. Esto nos lleva a un nuevo límite que reemplazará en forma adecuada al (14.45).

No exige gran esfuerzo tratar el caso más general de una sucesión de pruebas de Bernoulli, en la que la probabilidad de «éxito» o «suceso favorable» sea p y la de «fallo» o «suceso contrario» sea q. (En el lanzamiento de la moneda, el «éxito» será por ejemplo «cara» y tomamos $p = ½$.) Sea X la variable aleatoria que cuenta el número de sucesos favorables en n pruebas independientes. En tal caso X tiene distribución binomial con esperanza $E(X) = np$ y varianza $\text{Var}(X) = npq$. Luego la desigualdad de Chebyshev es aplicable; ella nos dice que

(14.46)
$$P(|X - np| > c) \leq \frac{npq}{c^2}.$$

Puesto que nos interesa la razón X/n, que podemos llamar *frecuencia relativa* del

suceso, dividimos la desigualdad $|X - np| > c$ por n y volvemos a escribir (14.46) así

$$P\left(\left|\frac{X}{n} - p\right| > \frac{c}{n}\right) \le \frac{npq}{c^2}. \tag{14.47}$$

Ya que esto es válido para todo $c > 0$, podemos hacer que c dependa de n y escribir $c = \epsilon n$, siendo ϵ un número positivo fijo. Entonces (14.47) se transforma en

$$P\left(\left|\frac{X}{n} - p\right| > \epsilon\right) \le \frac{pq}{n\epsilon^2}.$$

El hecho de que aparezca n en el denominador del segundo miembro sugiere que hagamos tender n a ∞. Esto nos conduce a la fórmula

$$\lim_{n\to\infty} P\left(\left|\frac{X}{n} - p\right| > \epsilon\right) = 0 \qquad \text{para todo } \epsilon > 0, \text{ fijo,} \tag{14.48}$$

que se denomina la *ley de los grandes números para la distribución de Bernoulli* Ella nos dice que, dado un $\epsilon > 0$ (por pequeño que sea), la probabilidad de que la frecuencia relativa del suceso difiera de p en más de ϵ es una función de n que tiende a 0 cuando $n \to \infty$. Esta relación nos da una justificación matemática para la asignación de la probabilidad ½ al suceso de conseguir cara con una moneda perfectamente equilibrada.

El límite (14.48) es un caso particular de un resultado más general en el que la «frecuencia relativa» X/n es reemplazada por la media aritmética de n variables aleatorias independientes con la misma esperanza y la misma varianza. Este teorema más general se conoce con el nombre de *ley débil de los grandes números;* puede establecerse así:

TEOREMA 14.12. LEY DÉBIL DE LOS GRANDES NÚMEROS. *Sean $X_1, X_2, \ldots, X_n$ n variables aleatorias independientes, todas con la misma esperanza y la misma varianza, y sean*

$$E(X_k) = m \qquad y \qquad \operatorname{Var}(X_k) = \sigma^2 \qquad para\ \ k = 1, 2, \ldots, n.$$

Definamos una nueva variable aleatoria $\bar{X}$ (llamada media aritmética de X_1, $X_2, \ldots, X_n$) mediante la igualdad

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

Entonces, para todo $\epsilon > 0$, fijo, tenemos

(14.49) $$\lim_{n\to\infty} P(|\bar{X} - m| > \epsilon) = 0 .$$

Una proposición equivalente es

(14.50) $$\lim_{n\to\infty} P(|\bar{X} - m| \leq \epsilon) = 1 .$$

Demostración. Aplicamos la desigualdad de Chebyshev a $\bar{X}$. Para ello necesitamos conocer la esperanza y la varianza de $\bar{X}$. Éstas son:

$$E(\bar{X}) = m \qquad \text{y} \qquad \operatorname{Var}(\bar{X}) = \frac{\sigma^2}{n} .$$

(Véase el ejercicio 5 de la sección 14.27.) La desigualdad de Chebyshev se convierte en $P(|\bar{X} - m| > c) \leq \sigma^2/(nc^2)$. Haciendo que $n \to \infty$ y sustituyendo c por ϵ obtenemos (14.49) y por lo tanto (14.50).

Observación. Para demostrar que el límite (14.48) es un caso particular del teorema 14.12, supongamos que cada X_k tiene los valores posibles 0 y 1, con probabilidades $P(X_k = 1) = p$ y $P(X_k = 0) = 1 - p$. Entonces $\bar{X}$ es la frecuencia relativa de éxito en n pruebas independientes, $E(\bar{X}) = p$, y (14.49) se reduce a (14.48).

El teorema 14.12 se denomina ley *débil* porque existe también una ley *fuerte* de los grandes números que (bajo las mismas hipótesis) establece que

(14.51) $$P\left(\lim_{n\to\infty} |\bar{X} - m| = 0\right) = 1 .$$

La diferencia esencial entre (14.51) y (14.50) es que las operaciones «límite» y «probabilidad» están intercambiadas. Puede demostrarse que la ley fuerte implica la débil, pero no al revés.

Obsérvese que la ley fuerte (14.51) parece ser más próxima a la (14.45) que a la (14.50). En realidad, (14.51) nos dice que $\lim_{n\to\infty} \bar{X} = m$ «casi siempre», es decir, con probabilidad 1. En particular, si la aplicamos al juego de cara y cruz, nos dice que el que no se cumpla la igualdad (14.45) es menos probable que el conseguir siempre cara al lanzar repetidamente una moneda correcta. La ley fuerte pone de manifiesto por qué la teoría de la probabilidad corresponde a la experiencia y a nuestra sensación intuitiva de lo que «debe ser» la probabilidad.

La demostración de la ley fuerte es larga y la omitiremos. Puede verse en las obras citadas en las referencias 1, 3, 8 y 10 del final de este capítulo.

14.30 El teorema central del límite

En muchas aplicaciones del Cálculo de probabilidades, las variables aleatorias son sumas de otras variables aleatorias. Por ejemplo, el resultado (en dinero) después de varias partidas de juego es la suma de las ganancias en cada una. Un hecho sorprendente acontece cuando se suman un gran número de variables aleatorias. Bajo condiciones generales (aplicable casi a todos los casos prácticos que se presentan) la distribución de la suma tiende a ser normal, prescindiendo de las distribuciones de cada una de las variables aleatorias que forman la suma. El enunciado preciso de este hecho notable se conoce como el *teorema central del límite del Cálculo de probabilidades*. Lo que explica la importancia de la distribución normal tanto en la teoría como en la práctica. La discusión completa de este teorema pertenece al estudio superior del Cálculo de probabilidades. En esta sección se explicará solamente lo que afirma el teorema.

Sea una sucesión indefinida de variables aleatorias, $X_1, X_2, \ldots$, con esperanzas y varianzas finitas. Sean éstas:

$$m_k = E(X_k) \qquad \text{y} \qquad \sigma_k^2 = \operatorname{Var}(X_k), \qquad k = 1, 2, \ldots.$$

Formamos una nueva variable aleatoria S_n sumando las n primeras diferencias $X_k - m_k$:

$$S_n = \sum_{k=1}^{n} (X_k - m_k). \tag{14.52}$$

En lugar de las X_k sumamos las *diferencias* de manera que la suma S_n tendrá por valor esperado el 0. El problema consiste en determinar la forma límite de la función de distribución de S_n cuando $n \to \infty$.

Si $X_1, X_2, \ldots, X_n$ son *independientes*, tenemos [en virtud del ejercicio 4 c) de la sección 14.27]

$$\operatorname{Var}(S_n) = \sum_{k=1}^{n} \operatorname{Var}(X_k - m_k) = \sum_{k=1}^{n} \operatorname{Var}(X_k) = \sum_{k=1}^{n} \sigma_k^2.$$

Ordinariamente, la $\operatorname{Var}(S_n)$ será grande aunque las varianzas particulares σ_k^2 sean pequeñas. Las variables aleatorias con gran varianza no interesan porque sus valores tienden a presentar gran dispersión respecto al valor esperado. Por este motivo, se introduce una nueva variable aleatoria T_n mediante la fórmula

$$T_n = \frac{S_n}{\sqrt{\operatorname{Var}(S_n)}}. \tag{14.53}$$

Esta variable tiene esperanza 0 y varianza 1 y se llama variable aleatoria *reducida*.

La variable aleatoria reducida T_n tiene pleno sentido aun cuando las variables $X_1, X_2, \ldots, X_n$ no sean independientes.

Introducimos ahora la siguiente definición:

DEFINICIÓN DE LA PROPIEDAD CENTRAL DEL LÍMITE. *Sea*

$$X_1, X_2, X_3, \ldots \tag{14.54}$$

una sucesión de variables aleatorias (no necesariamente independientes), donde cada X_k tiene una esperanza m_k y una varianza σ_k^2 ambas finitas. Se definen S_n y T_n mediante (14.52) *y* (14.53). *Se dice que la sucesión* (14.54) *satisface la propiedad central del límite si, para todo par a y b siendo $a \leq b$, tenemos*

$$\lim_{n \to \infty} P(a \leq T_n \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2}\, du\,. \tag{14.55}$$

Dicho de otro modo, las variables aleatorias (14.54) satisfacen la propiedad central del límite si la distribución de la variable reducida T_n se aproxima a una distribución normal cuando $n \to \infty$. [La igualdad (14.55) es válida también si $a = -\infty$ o $b = +\infty$.]

Laplace fue el primero en comprobar que muchas sucesiones de variables aleatorias satisfacen esta propiedad, si bien antes De Moivre conocía un caso particular (el de variables aleatorias que constituyen una sucesión de pruebas de Bernoulli). La figura 14.11 muestra una distribución binomial y la correspondiente aproximación normal. Laplace estableció un teorema central del límite general que fue demostrado por primera vez en forma completa por el matemático ruso A. Lyapunov en 1901. En 1922, J. W. Lindeberg generalizó el resultado de Laplace demostrando que la propiedad se cumple si las variables aleatorias son independientes y tienen una distribución común con las mismas esperanzas y varianzas, o sea $E(X_k) = m$ y $\operatorname{Var}(X_k) = \sigma^2$ para todo k. En tal caso la variable reducida es:

$$T_n = \frac{\sum_{k=1}^n X_k - nm}{\sigma\sqrt{n}}\,.$$

Lindeberg se dio cuenta de que tan sólo la independencia no es suficiente para garantizar la propiedad central del límite, pero formuló otra condición (la *condición de Lindeberg*) que, junto con la independencia *es* suficiente. En 1935, W. Feller demostró que la condición de Lindeberg es necesaria y suficiente para que las variables aleatorias independientes satisfagan la propiedad central del límite. Aquí no haremos la discusión de la condición de Lindeberg, sólo mencionaremos que ella implica

$$\operatorname{Var}(S_n) \to \infty \text{ cuando } n \to \infty\,.$$

Afortunadamente, muchas variables aleatorias independientes que en la práctica se presentan, satisfacen la condición de Lindeberg y por consiguiente también la propiedad central del límite. Hasta ahora, la teoría para variables aleatorias *dependientes* es incompleta. Tan sólo se han tratado unos pocos casos particulares. La investigación contemporánea en Cálculo de probabilidades en gran parte está centrada en la búsqueda de teoremas generales referentes a variables dependientes.

14.31 Ejercicios

1. Demostrar la desigualdad de Chebyshev en el caso discreto.
2. Si a es un número real cualquiera, demostrar que

$$P(|X - a| > c\lambda) \leq \frac{1}{c^2}$$

para todo $c > 0$, siendo $\lambda^2 = \int_{-\infty}^{+\infty} (t - a)^2 f_X(t)\, dt$. La desigualdad de Chebyshev es el caso particular en el que $a = E(X)$.

3. Sea X la variable aleatoria que cuenta el número de éxitos en n pruebas independientes de una sucesión de Bernoulli; la probabilidad de éxito es p. Demostrar que, para todo $\epsilon > 0$,

$$P\left(\left|\frac{X}{n} - p\right| > \epsilon\right) \leq \frac{1}{4n\epsilon^2}.$$

4. Se lanza n veces una moneda «correcta»; el número de caras se representa por X. Encontrar el menor valor de n para el que la desigualdad de Chebyshev implique

$$P\left(0.4 < \frac{X}{n} < 0.6\right) > 0.90.$$

5. En una cadena de producción el número X de artículos manufacturados defectuosos en una determinada hora se sabe que tiene una distribución de Poisson con media $E(X) = 100$. Emplear la desigualdad de Chebyshev para calcular una cota inferior para la probabilidad de que en una hora determinada los artículos defectuosos producidos esté comprendidos entre 90 y 110.
6. Supongamos que una variable aleatoria X tenga una distribución normal canónica (media 0 y varianza 1). Representemos con p la probabilidad de que X difiera de su esperanza $E(X)$ en más de tres veces la desviación típica. Aplicar la desigualdad de Chebyshev para encontrar una cota superior de p. Seguidamente, usando las correspondientes tablas de la distribución normal, comprobar que hay una cota superior de p que es aproximadamente igual a un cincuentavo de la obtenida por la desigualdad de Chebyshev.
7. Dada una sucesión de variables independientes $X_1, X_2, \ldots$, con distribución normal cada una. Sean $m_k = E(X_k)$ y $\sigma_k^2 = \text{Var}(X_k)$. Probar que tal sucesión tiene la propiedad central del límite. [*Indicación:* Recuérdese el ejercicio 7 de la sección 14.24.]
8. Sean las variables aleatorias independientes $X_1, X_2, \ldots$, con la misma distribución binomial. Supongamos que cada X_k toma los valores 0 y 1 con probabilidades $P(X_k = 1) = p$ y $P(X_k = 0) = q$, siendo $p + q = 1$. La variable aleatoria $Z_n = X_1 + X_2 + \ldots + X_n$ cuenta el número de éxitos en n pruebas de Bernoulli.

a) Demostrar que la propiedad central del límite toma la forma siguiente:

$$\lim_{n\to\infty} P\left(\frac{Z_n - np}{\sqrt{npq}} \leq t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2}\, du .$$

b) Usar la aproximación sugerida en la parte a) para estimar la probabilidad de obtener un número de caras comprendido entre 45 y 55 si se lanza 100 veces una moneda correcta. Recurrir a tablas de distribución normal para realizar el cálculo.

9. Con la notación del ejercicio 8, el teorema central del límite para variables aleatorias que forman una sucesión de pruebas de Bernoulli puede escribirse en la forma

$$\lim_{n\to\infty} \frac{P\left(t_1 \leq \dfrac{Z_n - np}{\sqrt{npq}} \leq t_2\right)}{\Phi(t_2) - \Phi(t_1)} = 1 ,$$

donde Φ es la distribución normal canónica. En este caso particular puede demostrarse que la fórmula también es válida cuando t_1 y t_2 son funciones de n dadas por $t_1 = (a - np)/\sqrt{npq}$ y $t_2 = (b - np)/\sqrt{npq}$, donde a y b son constantes positivas prefijadas, $a < b$.

a) Demostrar que esta relación implica la fórmula asintótica

$$\sum_{k=a}^{b} \binom{n}{k} p^k q^{n-k} \sim \Phi\left(\frac{b - np + \frac{1}{2}}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np - \frac{1}{2}}{\sqrt{npq}}\right) \text{ cuando } n \to \infty .$$

b) Se lanza un dado insesgado o correcto 180 veces. Usar la aproximación sugerida en la parte a) para estimar la probabilidad de que salga un seis exactamente 30 veces. Para los cálculos hacer uso de las tablas de distribución normal 14.1.

10. Se lanza un dado insesgado o correcto 100 veces. Utilizar la aproximación sugerida en el ejercicio 9 a) para estimar la probabilidad de que salga un seis a) exactamente 25 veces, b) por lo menos 25 veces. Para los cálculos utilizar tablas de distribución normal.

Referencias citadas

1. H. Cramér, *Elements of Probability Theory,* John Wiley, New York, 1955.
2. H. Cramér, *Mathematical Methods of Statistics,* Princeton Univ. Press, Princeton, N. J., 1946.
3. W. Feller, *An Introduction to Probability Theory and its Applications,* 2.ª edición John Wiley, New York, 1957.
4. B. V. Gnedenko y A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables,* Addison-Wesley, Reading, Mass., 1954.
5. S. Goldberg, *Probability, an Introduction,* Prentice-Hall, Englewood Cliffs, N. J., 1960.
6. H. Levy y L. Roth, *Elements of Probability,* Oxford Univ. Press, London y New York, 1936.
7. M. Loève, *Probability Theory: Foundations, Random Sequences,* Van Nostrand, New York, 1955.

8. M. E. Munroe, *Theory of Probability*, McGraw-Hill Book Co., Inc., New York, 1951.
9. J. Neyman, *First Course in Probability and Statistics*, Holt, Rinehart and Winston, New York, 1950.
10. E. Parzen, *Modern Probability Theory and its Applications*, John Wiley, New York, 1960.
11. I. Todhunter, *A History of the Mathematical Theory of Probability from the Time of Pascal to Laplace*, Chelsea Publishing Co., New York, 1949.
12. J. V. Uspensky, *Introduction to Mathematical Probability*, MacGraw-Hill Book Co., Inc., New York, 1937.

15

INTRODUCCIÓN AL ANÁLISIS NUMÉRICO

15.1 Introducción histórica

El planeta Urano fue descubierto en 1781 por un inteligente astrónomo aficionado, William Herschel (1738-1822), con un telescopio de fabricación casera. Teniendo en cuenta las leyes de Kepler, la prevista órbita de Urano fue rápidamente calculada a partir de unas pocas observaciones muy separadas entre sí. Se encontró que la distancia media de Urano al Sol era aproximadamente el doble que la de Saturno y que una órbita completa requeriría 84 años. En 1830 los datos empíricos acumulados pusieron de manifiesto desviaciones inexplicables de la órbita prevista. Algunos astrónomos llegaron a pensar que la ley de la gravitación universal de Newton no fuera válida para distancias tan grandes como la de Urano al Sol; otros sospecharon que las perturbaciones fueran debidas a un cometa aún no descubierto o a un planeta más lejano.

Un estudiante del bachillerato de la Universidad de Cambridge, John Couch Adams (1819-1892), estaba intrigado por la posibilidad de un planeta desconocido. Se asignó la difícil tarea de calcular la influencia de un tal planeta en las posiciones observadas de Urano, suponiendo válida la ley de gravitación de Newton. Completó su cálculo en 1845 e instó al Real Observatorio de Greenwich a buscar el supuesto planeta, pero su requerimiento no fue tomado en serio.

Independientemente y casi simultáneamente Jean Joseph Leverrier (1811-1877) de París, realizó un cálculo parecido y pidió a Johann Galle, jefe del Observatorio de Berlín, que confirmara su predicción. La misma noche que recibió la carta de Leverrier, Galle encontró el nuevo planeta, Neptuno, casi exactamente en la posición calculada. Éste fue otro triunfo de la ley de gravitación de Newton, y uno de los primeros grandes triunfos del *análisis numérico,* el arte y la ciencia de calcular.

La historia del análisis numérico data de tiempos antiguos. Los babilonios, 2000 años a.J., compusieron tablas matemáticas. Se ha encontrado una tablilla

de barro con los cuadrados de los enteros del 1 al 60. Los babilonios adoraban los cuerpos celestes y elaboraban efemérides astronómicas. El famoso astrónomo alejandrino Claudio Ptolomeo (aproximadamente 150 d.J.) poseía unas efemérides babilónicas de eclipses que databan del año 747 a.J.

Arquímedes, en el año 220 a.J., usó los polígonos regulares como aproximaciones del círculo y dedujo las desigualdades $3\frac{10}{71} < \pi < 3\frac{1}{7}$. El trabajo de cálculo numérico desde entonces hasta el siglo XVII fue centrado principalmente en la preparación de tablas astronómicas. El advenimiento del Álgebra en el siglo XVI produjo una renovada actividad en todas las ramas de la Matemática, incluyendo el análisis numérico. En 1614, Neper publicó la primera tabla de logaritmos, y en 1620, los logaritmos de las funciones seno y tangente fueron tabuladas con siete cifras decimales. Hacia 1628 habían sido calculadas tablas de logaritmos con catorce decimales de los números 1 al 100 000.

El cálculo con series empezó a florecer hacia fines del siglo XVII, con el desarrollo del cálculo. A principios del siglo XVIII Jacob Stirling y Brook Taylor sentaron los fundamentos del *cálculo de diferencias finitas,* que ahora desempeña un papel central en el análisis numérico. Con la predicción de la existencia y la localización del planeta Neptuno por Adam y Leverrier en 1845, la importancia científica del análisis numérico quedó establecida de una vez para siempre.

A fines del siglo XIX, el empleo de las máquinas de cálculo automático estimuló aún más el desarrollo del análisis numérico. Tal desarrollo ha sido explosivo desde la terminación de la segunda Guerra Mundial a causa del progreso en las máquinas de cálculo electrónicas de alta velocidad. Las nuevas máquinas han hecho posibles gran número de importantes logros científicos que antes parecían inaccesibles.

El arte de calcular (distinto de la ciencia de calcular) da gran importancia a la elaboración minuciosa del plan que se necesita en un determinado cálculo. También trata cuestiones tales como, la precisión y exactitud, los errores, y la comprobación. Este aspecto del análisis numérico no lo vamos aquí a discutir; se aprende mejor al efectuar los cálculos numéricos en problemas concretos. Para tener un buen conocimiento de los métodos prácticos y técnicos a seguir, el lector debería consultar los libros que existen sobre análisis numérico, algunos de los cuales se citan en la Bibliografía al final de este capítulo. La bibliografía contiene también algunas de las tablas matemáticas más usadas, muchas de las cuales dan también información práctica de cómo hay que proceder en un cálculo determinado.

Este capítulo es una introducción a la *ciencia* del cálculo. Contiene algunos de los principios matemáticos básicos que precisa quien maneja el análisis numérico, tanto si trabaja en gabinete o con una máquina calculadora de alta velocidad. Aparte de su valor práctico, la materia de este capítulo tiene interés por derecho propio, y es de esperar que esta breve introducción estimulará al lector a profundizar más en esta importante y fascinante rama de la Matemática.

15.2 Aproximaciones por polinomios

Una idea básica en análisis numérico es la de utilizar funciones sencillas, de ordinario polinomios, para aproximar una función dada f. En el volumen I se vio un tipo de polinomio de aproximación al tratar de la fórmula de Taylor (teorema 7.1). El problema fue entonces encontrar un polinomio P que coincidiera con una función dada f y con algunas de sus derivadas en un punto dado. Se demostró que si f es una función con derivada de orden n en un punto a, existe un polinomio P y sólo uno, de grado $\leq n$ que satisface las $n + 1$ relaciones

$$P(a) = f(a), \qquad P'(a) = f'(a), \qquad \ldots, \qquad P^{(n)}(a) = f^{(n)}(a).$$

La solución viene dada por el *polinomio de Taylor,*

$$P(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k.$$

También se discutió el error que se comete aproximando $f(x)$ por $P(x)$ en puntos x distintos del a. Tal error se define por la diferencia $E_n(x) = f(x) - P(x)$, de modo que podemos escribir

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k + E_n(x).$$

Para estudiar más profundamente el error necesitamos mayor información acerca de la función f. Por ejemplo, si f posee derivada de orden $n + 1$ continua en un cierto intervalo que contenga a, entonces para todo x de ese intervalo el error puede expresarse en forma de integral o como una derivada de orden $n + 1$:

$$E_n(x) = \frac{1}{n!}\int_a^x (x - t)^n f^{(n+1)}(t)\, dt = \frac{f^{(n+1)}(c)}{(n + 1)!}(x - a)^{n+1},$$

siendo c un punto situado entre a y x. (Ver secciones 7.5 y 7.7 del volumen I.)

Existen muchos otros métodos para aproximar una función dada f mediante polinomios, dependiendo del uso que debe hacerse de la aproximación. Por ejemplo, en lugar de querer hallar un polinomio que coincida con f y con algunas de sus derivadas en un punto dado, podemos desear un polinomio que tome los mismos valores que f en un cierto número de puntos distintos. Concretamente, si

los puntos dados son $x_0, x_1, \ldots, x_n$ podemos buscar un polinomio P que satisfaga las condiciones

$$(15.1) \qquad P(x_0) = f(x_0), \qquad P(x_1) = f(x_1), \qquad \ldots, \qquad P(x_n) = f(x_n).$$

Puesto que hay $n + 1$ condiciones a satisfacer intentamos resolver el problema con un polinomio de grado $\leq n$, sea éste

$$P(x) = \sum_{k=0}^{n} a_k x^k,$$

con $n + 1$ coeficientes $a_0, a_1, \ldots, a_n$ a determinar. Las $n + 1$ condiciones (15.1) nos conducen a un sistema de $n + 1$ ecuaciones de primer grado entre los coeficientes. Según la teoría de las ecuaciones lineales puede demostrarse que este sistema tiene una y una sola solución; luego el citado polinomio existe siempre. Si el sistema se resuelve por la regla de Cramer, los coeficientes $a_0, a_1, \ldots, a_n$ se expresan como cocientes de determinantes. En la práctica, no obstante, el polinomio P rara vez se determina de este modo debido a que los cálculos son extremadamente laboriosos cuando n es grande. Se han desarrollado métodos más sencillos para calcular los polinomios de aproximación. Algunos de ellos serán discutidos en secciones posteriores. El polinomio que resuelve el problema antes citado se llama *polinomio de interpolación.*

Otro tipo corriente de polinomio de aproximación es el llamado de *aproximación por mínimos cuadrados.* En este caso, la función dada f está definida y es integrable en un intervalo $[a, b]$ y buscamos un polinomio P de grado $\leq n$ tal que el error cuadrático medio

$$\int_a^b |f(x) - P(x)|^2 \, dx$$

sea lo menor posible. En la sección 15.4 demostraremos que para una función continua f tal polinomio existe y es único. Los polinomios de Legendre introducidos en la sección 1.14 desempeñan un papel fundamental en la resolución de ese problema.

15.3 Aproximaciones polinómicas y espacios lineales normados

Los distintos tipos de aproximación por polinomios descritos en la sección anterior están relacionados por medio de una idea central que se describe mejor con la nomenclatura de los espacios lineales.

Sea V un espacio lineal de funciones que contiene todos los polinomios de grado $\leq n$ y que también contiene la función f que debe aproximarse. Los polinomios constituyen un subespacio S de dimensión finita, siendo dim $S = n + 1$.

Cuando hablamos de la aproximación de f mediante un polinomio P de S, consideramos la diferencia $f - P$, que llamamos error de la aproximación, y tratamos entonces de encontrar un método para medir la magnitud de ese error.

Si V es un espacio euclídeo, tiene un producto interior (x, y) y una correspondiente norma dada por $\|x\| = (x, x)^{1/2}$, y podemos usar la norma $\|f - P\|$ como medida del tamaño del error.

Algunas veces pueden introducirse normas en espacios lineales no euclídeos, esto es, en espacios lineales desprovistos de producto interior. Tales normas se introdujeron en la sección 7.26. Repetimos aquí la definición.

DEFINICIÓN DE NORMA. *Sea V un espacio lineal de funciones. Una función real N definida en V se llama norma si tiene las propiedades siguientes:*

a) $N(f) \geq 0$ *para cada* f *de* V.

b) $N(cf) = |c|N(f)$ *para cada* f *de* V *y cada número real* c.

c) $N(f + g) \leq N(f) + N(g)$ *para todo par de funciones* f *y* g *de* V.

d) $N(f) = 0$ *implica* $f = 0$.

Un espacio lineal con una norma asignada se denomina espacio lineal normado.

Algunas veces la norma de f se representa por $\|f\|$ en lugar de $N(f)$. Con esta notación, las propiedades fundamentales se escriben:

a) $\|f\| \geq 0$,

b) $\|cf\| = |c|\,\|f\|$,

c) $\|f + g\| \leq \|f\| + \|g\|$,

d) $\|f\| = 0$ implica $f = 0$.

Una función N que satisface las propiedades a), b) y c), pero *no* la d), se llama *seminorma*. Algunos problemas de la teoría de la aproximación tratan con espacios lineales seminormados; otros con espacios lineales normados. Los ejemplos siguientes se discutirán en este capítulo.

EJEMPLO 1. *Seminorma de Taylor.* Para un entero fijo $n \geq 1$, sea V el conjunto de todas las funciones que poseen derivada de orden n en un punto fijo a. El conjunto V es un espacio lineal. Si $f \in V$, sea

$$N(f) = \sum_{k=0}^{n} |f^{(k)}(a)|\,.$$

Es fácil comprobar que la función N así definida es una seminorma. No es una norma porque $N(f) = 0$ si y sólo si

$$f(a) = f'(a) = \cdots = f^{(n)}(a) = 0\,,$$

y esas ecuaciones pueden satisfacerse con una función no nula. Por ejemplo, $N(f) = 0$ cuando $f(x) = (x - a)^{n+1}$.

EJEMPLO 2. *Seminorma de interpolación.* Sea V el espacio lineal de todas las funciones reales definidas en un intervalo $[a, b]$. Para un conjunto fijo de $n + 1$ puntos distintos $x_0, x_1, \ldots, x_n$ de $[a,b]$, definamos N mediante la ecuación

$$N(f) = \sum_{k=0}^{n} |f(x_k)|$$

si $f \in V$. Esta función N es una seminorma en V. No es una norma porque $N(f) = 0$ si y sólo si $f(x_0) = f(x_1) = \ldots = f(x_n) = 0$, y es evidente que esas ecuaciones se pueden satisfacer con una función f no nula en $[a, b]$.

EJEMPLO 3. *Norma cuadrática.* Sea C el espacio lineal de las funciones continuas en un intervalo $[a, b]$. Si $f \in C$ definimos

$$N(f) = \left(\int_a^b |f(x)|^2\, dx\right)^{1/2}. \tag{15.2}$$

Esta es una norma consecuencia del producto interior.

$$(f, g) = \int_a^b f(x)\overline{g(x)}\, dx.$$

Observación. Representemos con S el conjunto de todas las funciones f que son integrables en $[a, b]$. El conjunto S es un espacio lineal, y la función N definida en (15.2) es una seminorma en S. No es una norma porque puede ser $N(f)=0$ sin que f sea idénticamente nula en $[a, b]$.

EJEMPLO 4. *Norma del máximo.* Representemos con C el espacio lineal de las funciones continuas en un intervalo $[a, b]$. Si $f \in C$, definimos

$$N(f) = \max_{a \le x \le b} |f(x)|,$$

en donde el símbolo del segundo miembro representa el máximo absoluto de $|f|$ en $[a, b]$. En el ejercicio 4 de la sección 15.5 se pide la comprobación de las cuatro propiedades de la norma.

15.4 Problemas fundamentales en la aproximación por polinomios

Sea C el espacio de las funciones continuas en un intervalo dado $[a, b]$, y sea S un subespacio lineal que contiene todos los polinomios de grado $\le n$. Supongamos también que en C se ha definido una norma o una seminorma. Elijamos una función f de C. Si existe un polinomio P en S tal que

$$\|f - P\| \le \|f - Q\|$$

para todo polinomio Q de S, decimos que P es el *polinomio de aproximación óptima* para f con grado asignado. La palabra «óptima» es relativa a la norma dada (o la seminorma). El polinomio de aproximación óptima para una cierta norma elegida no lo es necesariamente si se elige otra.

Una vez elegidas una norma o seminorma, se presentan inmediatamente tres problemas.

1. *Existencia.* Dada f de C, ¿existe el polinomio de aproximación óptima para f con grado prefijado?

2. *Unicidad.* Si existe ese polinomio, ¿es único?

3. *Construcción.* ¿Cómo puede determinarse, si existe, el polinomio de aproximación óptima para f con grado prefijado?

Existen, como es natural, otros muchos problemas que pueden considerarse. Por ejemplo, si existe un único polinomio de aproximación óptima P_n de grado $\leq n$, podemos desear obtener cotas superiores de $\|f - P_n\|$ que se puedan utilizar para satisfacer ciertas necesidades prácticas. O podemos averiguar si $\|f - P_n\| \to 0$ cuando $n \to \infty$ para una norma dada o para alguna otra norma. Si es así, decimos que los polinomios de aproximación convergen hacia f en esta norma. En tal caso existen aproximaciones tan ajustadas como se quiera con relación a esa norma si n es suficientemente grande. Estos ejemplos muestran algunos de los problemas que se estudian en la teoría general de la aproximación con polinomios. En este estudio preliminar dedicamos nuestra atención especialmente a los tres problemas de existencia, unicidad y construcción antes citados.

Esos tres problemas se pueden resolver totalmente en el caso de la aproximación mediante polinomios de Taylor. Si f posee derivada de orden n en un punto a, es fácil probar que el polinomio de aproximación óptima de grado $\leq n$ relativo a la seminorma de Taylor para este n es el polinomio de Taylor

$$P(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k. \tag{15.3}$$

En efecto, para este polinomio tenemos

$$\|f - P\| = \sum_{k=0}^{n} |f^{(k)}(a) - P^{(k)}(a)| = 0,$$

así que la desigualdad $\|f - P\| \leq \|f - Q\|$ se satisface para todo polinomio Q. Por lo tanto, P es el polinomio de aproximación óptima relativo a esa seminorma.

Para establecer la unicidad, consideramos un polinomio cualquiera Q de grado $\leq n$ tal que $\|f - Q\| = 0$. Esta ecuación implica que

$$Q(a) = f(a), \qquad Q'(a) = f'(a), \qquad \ldots, \qquad Q^{(n)}(a) = f^{(n)}(a).$$

Según el teorema 7.1 del volumen I sabemos que el polinomio de Taylor (15.3) es el único polinomio que satisface todas esas condiciones. Por consiguiente, $Q = P$. La ecuación (15.3) resuelve también el problema de la construcción.

Los tres problemas pueden también resolverse para cualquier norma deducida de un producto interior. En este caso, el teorema 1.16 nos dice que existe un solo polinomio en S para el cual la norma $\|f - P\|$ es tan pequeña como se quiera. En efecto, este P es la proyección de f sobre S y viene dado por la fórmula explícita,

$$P(x) = \sum_{k=0}^{n} (P, e_i)e_i(x),$$

en donde $e_0, e_1, \ldots, e_n$ son funciones que constituyen una base ortonormal para S.

Por ejemplo, si C es el espacio de las funciones reales continuas en el intervalo $[-1, 1]$ y si

$$(f, g) = \int_{-1}^{1} f(x)g(x)\, dx,$$

los polinomios de Legendre normalizados $\varphi_0, \varphi_1, \ldots, \varphi_n$ forman una base ortonormal para S, y la proyección f_n de f sobre S viene dada por

$$f_n(x) = \sum_{k=0}^{n} (f, \varphi_k)\varphi_k(x), \qquad \text{donde} \quad (f, \varphi_k) = \int_{-1}^{1} f(t)\varphi_k(t)\, dt.$$

Recordemos que los polinomios de Legendre normalizados vienen dados por

$$\varphi_k(x) = \sqrt{\frac{2k+1}{2}}\, P_k(x), \qquad \text{donde} \quad P_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k}(x^2 - 1)^k.$$

Los seis primeros son

$$\varphi_0(x) = \sqrt{\tfrac{1}{2}}, \quad \varphi_1(x) = \sqrt{\tfrac{3}{2}}\, x, \quad \varphi_2(x) = \tfrac{1}{2}\sqrt{\tfrac{5}{2}}\,(3x^2 - 1), \quad \varphi_3(x) = \tfrac{1}{2}\sqrt{\tfrac{7}{2}}\,(5x^3 - 3x),$$

$$\varphi_4(x) = \tfrac{1}{8}\sqrt{\tfrac{9}{2}}\,(35x^4 - 30x^2 + 3), \qquad \varphi_5(x) = \tfrac{1}{8}\sqrt{\tfrac{11}{2}}\,(63x^5 - 70x^3 + 15x).$$

Los problemas correspondientes a la seminorma de interpolación se tratarán

en la próxima sección 15.6. En secciones posteriores estudiaremos la aproximación por polinomios relativa a la norma del máximo.

15.5 Ejercicios

1. Demostrar que cada uno de los siguientes conjuntos de funciones es un espacio lineal
   a) Todos los polinomios.
   b) Todos los polinomios de grado $\leq n$.
   c) Todas las funciones continuas en un intervalo I.
   d) Todas las funciones derivables en cada punto de I.
   e) Todas las funciones que tienen derivada de orden n en cada punto de I.
   f) Todas las funciones con derivadas de orden n en un punto fijo x_0.
   g) Todas las funciones desarrollables en serie de potencias en un entorno de un punto dado x_0.
2. Determinar si cada uno de los siguientes conjuntos de funciones es o no un espacio lineal.
   a) Todos los polinomios de grado n.
   b) Todas las funciones definidas y acotadas en un intervalo $[a, b]$.
   c) Todas las funciones definidas en un intervalo $[a, b]$.
   d) Todas las funciones monótonas en un intervalo $[a, b]$.
   e) Todas las funciones integrables en un intervalo $[a, b]$.
   f) Todas las funciones regulares a trozos monótonas en un intervalo $[a, b]$.
   g) Todas las funciones que pueden expresarse en la forma $f - g$, siendo f y g monótonas crecientes en un intervalo $[a, b]$.
3. Sea C el espacio lineal de las funciones reales continuas en un intervalo $[a, b]$. Mediante la ecuación que se da, se define en C una función N. En cada caso, determinar cuáles de las cuatro propiedades de una norma satisface N, y determinar de este modo si N es una norma, una seminorma, o ni una cosa ni otra.

   a) $N(f) = f(a)$.
   b) $N(f) = |f(a)|$.
   c) $N(f) = |f(b) - f(a)|$.
   d) $N(f) = \int_a^b f(x)\, dx$.
   e) $N(f) = \left| \int_a^b f(x)\, dx \right|$.
   f) $N(f) = \int_a^b |f(x)|\, dx$.
   g) $N(f) = \int_a^b |f(x)|^2\, dx$.
   h) $N(f) = \left| \int_a^b f(x)\, dx \right|^2$.

4. Sea C el espacio lineal de las funciones continuas en un intervalo $[a, b]$. Si $f \in C$, definimos

$$N(f) = \max_{a \leq x \leq b} |f(x)|.$$

   Demostrar que N es una norma para C.
5. Representemos por B el espacio lineal de todas las funciones reales definidas y acotadas en un intervalo $[a, b]$. Si $f \in B$ definimos

$$N(f) = \sup_{a \leq x \leq b} |f(x)|,$$

donde el símbolo del segundo miembro representa el extremo superior del conjunto de todos los números $|f(x)|$ para x en $[a, b]$. Demostrar que N es una norma para B. Esta es la *norma del supremo.*

6. Refiriéndose al ejercicio 3, determinar cuáles de las funciones dadas N tienen la propiedad de que $N(fg) \leq N(f)N(g)$ para todo par de funciones f y g de C.
7. Para un entero fijo $n \geq 1$, sea S el conjunto de todas las funciones con derivada n-ésima en un punto x_0. Si $f \in S$, sea

$$N(f) = \sum_{k=0}^{n} \frac{1}{k!} |f^{(k)}(x_0)| .$$

a) Demostrar que N es una seminorma en S.
b) Demostrar que $N(fg) \leq N(f)N(g)$ cualesquiera que sean f y g de S. Demostrar también que la seminorma de Taylor no goza de esa propiedad.

8. Sea f una función real continua en el intervalo $[-1, 1]$.
a) Demostrar que el polinomio cuadrático de aproximación óptima relativo a la norma cuadrática en $[-1, 1]$ viene dado por

$$P(x) = \tfrac{1}{2} \int_{-1}^{1} f(t)\,dt + \tfrac{3}{2}x \int_{-1}^{1} tf(t)\,dt + \tfrac{5}{8}(3x^2 - 1) \int_{-1}^{1} (3t^2 - 1)f(t)\,dt .$$

b) Encontrar una fórmula análoga para el polinomio de aproximación óptima de grado ≤ 4.

9. Calcular las constantes a, b, c de modo que la integral $\int_{-1}^{1} |e^x - (a + bx + cx^2)|^2\,dx$ llegue a ser tan pequeña como se quiera.
10. Sea $f(x) = |x|$ para $-1 \leq x \leq 1$. Determinar el polinomio de grado ≤ 4 de aproximación óptima para f en $[-1, 1]$ relativo a la norma cuadrática.
11. Sea C el espacio lineal de las funciones reales continuas en $[a, b]$ con producto interior $(f, g) = \int_a^b f(x)g(x)\,dx$. Sea $e_0, \ldots, e_n$ una base ortonormal para el subespacio S de los polinomios de grado $\leq n$. Sea P el polinomio de S de aproximación óptima para f en C relativo a la norma cuadrática.
a) Demostrar que el cuadrado de la norma del error viene dado por

$$\|f - P\|^2 = \|f\|^2 - \sum_{k=0}^{n} (f, e_k)^2 .$$

b) Calcular explícitamente ese error cuando $[a, b] = [-1, 1]$, $n = 2$, y $f(x) = |x|$.

12. Sea $f(x) = 1/x$ para $x \neq 0$.
a) Demostrar que el polinomio constante P de aproximación óptima para f en el intervalo $[1, n]$ relativo a la norma cuadrática es $P(x) = (\log n)/(n - 1)$. Calcular $\|P - f\|^2$ para este P.
b) Hallar el polinomio lineal P de aproximación óptima para f en el intervalo $[1, n]$ relativo a la norma cuadrática. Calcular $\|P - f\|^2$ para este P cuando $n = 2$.

13. Sea $f(x) = e^x$.
a) Demostrar que el polinomio constante P de aproximación óptima para f en el intervalo $[0, n]$ relativo a la norma cuadrática es $P(x) = (e^n - 1)/n$. Calcular $\|P - f\|^2$ para este P.

b) Hallar el polinomio lineal P de aproximación óptima para f en el intervalo $[0, 1]$ relativo a la norma cuadrática. Calcular $\|P - f\|^2$ para este P.

14. Sean $P_0, P_1, \ldots, P_n$ $n + 1$ polinomios ortonormales en $[a, b]$ relativos al producto interior del ejercicio 11. Supóngase también que P_k es de grado k.

a) Demostrar que tres polinomios consecutivos cualesquiera de ese conjunto están ligados por una relación de recurrencia de la forma

$$P_{k+1}(x) = (a_k x + b_k)P_k(x) + c_k P_{k-1}(x)$$

para $1 \le k \le n - 1$, siendo a_k, b_k, c_k constantes.

b) Determinar explícitamente esa relación de recurrencia cuando los polinomios son los polinomios de Legendre ortonormales.

15. En relación con el ejercicio 14, y designando por p_k los coeficientes de x^k en $P_k(x)$.

a) Demostrar que $a_k = p_{k+1}/p_k$.

b) Utilizar la relación de recurrencia del ejercicio 14 para deducir la fórmula

$$\sum_{k=0}^{m} P_k(x)P_k(y) = \frac{p_m}{p_{m+1}} \frac{P_{m+1}(x)P_m(y) - P_m(x)P_{m+1}(y)}{x - y},$$

válida para $x \neq y$. Discutir también el caso límite $x = y$.

15.6 Polinomios de interpolación

Volvemos ahora a la aproximación mediante polinomios de interpolación. Se conocen los valores de una función f en $n + 1$ puntos distintos $x_0, x_1, \ldots, x_n$ y buscamos un polinomio P de grado $\le n$ que satisfaga las condiciones

$$(15.4) \qquad P(x_0) = f(x_0), \qquad P(x_1) = f(x_1), \qquad \ldots, \qquad P(x_n) = f(x_n).$$

Primero demostramos que si existe tal polinomio, éste es único. Luego probamos su existencia por medio de su construcción explícita. Este polinomio hace mínima la distancia de f a P, medida con la seminorma de interpolación para ese n,

$$\|f - P\| = \sum_{k=0}^{n} |f(x_k) - P(x_k)|.$$

Puesto que la distancia es 0 si P satisface (15.4), el polinomio de interpolación P es el de aproximación óptima relativo a aquella seminorma.

TEOREMA 15.1. TEOREMA DE UNICIDAD. *Dados $n + 1$ puntos distintos x_0, $x_1, \ldots, x_n$, sean P y Q dos polinomios de grado $\le n$ tales que*

$$P(x_k) = Q(x_k)$$

para cada $k = 0, 1, 2, \ldots, n$. Entonces $P(x) = Q(x)$ para todo x.

Demostración. Sea $R(x) = P(x) - Q(x)$. La función R es un polinomio de grado $\leq n$ que tiene $n + 1$ ceros distintos en los puntos $x_0, x_1, \ldots, x_n$. El único polinomio que goza de esta propiedad es el polinomio nulo. Por lo tanto, $R(x) = 0$ para todo x, luego $P(x) = Q(x)$ para todo x.

El polinomio de interpolación P puede construirse de varias maneras. Exponemos primero un método de Lagrange. Sea $A(x)$ el polinomio dado por la ecuación

(15.5) $$A(x) = (x - x_0)(x - x_1) \cdots (x - x_n) = \prod_{j=0}^{n} (x - x_j).$$

Este polinomio tiene un cero simple en cada uno de los puntos x_j. Sea $A_k(x)$ el polinomio de grado n obtenido del $A(x)$ por la supresión del factor $x - x_k$. Esto es,

(15.6) $$A_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} (x - x_j).$$

El polinomio $A_k(x)$ tiene un cero simple en cada punto $x_j \neq x_k$. En el punto x_k tenemos

(15.7) $$A_k(x_k) = \prod_{\substack{j=0 \\ j \neq k}}^{n} (x_k - x_j).$$

Este producto no es cero ya que ningún factor lo es. Por consiguiente, el polinomio $A_k(x)/A_k(x_k)$ toma el valor 1 cuando $x = x_k$ y el valor 0 cuando $x = x_j$ para $x_j \neq x_k$. Sea ahora

$$P(x) = \sum_{k=0}^{n} \frac{f(x_k)A_k(x)}{A_k(x_k)}.$$

Cuando $x = x_j$, cada término de esta suma se anula excepto el término de lugar j, que toma el valor $f(x_j)$. Por consiguiente, $P(x_j) = f(x_j)$ para cada j. Puesto que cada término de esta suma es un polinomio de grado n, la suma es a su vez un polinomio de grado $\leq n$. Así pues, hemos encontrado un polinomio que satisface las condiciones exigidas. Estos resultados podemos resumirlos en el teorema siguiente:

TEOREMA 15.2. *Dados $n + 1$ puntos distintos $x_0, x_1, \ldots, x_n$ y $n + 1$ números reales $f(x_0), f(x_1), \ldots, f(x_n)$, no necesariamente distintos, existe un polino-*

mio P y sólo uno de grado $\leq n$ *tal que* $P(x_j) = f(x_j)$ *para cada* $j = 0, 1, 2, \ldots, n$. *Este polinomio viene dado por la fórmula*

$$(15.8) \qquad P(x) = \sum_{k=0}^{n} \frac{f(x_k)A_k(x)}{A_k(x_k)},$$

en donde $A_k(x)$ *es el polinomio definido por* (15.6).

La fórmula (15.8) que da $P(x)$ se llama *fórmula de interpolación de Lagrange.* Podemos ponerla en la forma

$$P(x) = \sum_{k=0}^{n} f(x_k)L_k(x),$$

donde $L_k(x)$ es un polinomio de grado n dado por

$$(15.9) \qquad L_k(x) = \frac{A_k(x)}{A_k(x_k)}.$$

De este modo, para cada x fijo, $P(x)$ es una combinación lineal de los valores prefijados $f(x_0), f(x_1), \ldots, f(x_n)$. Los multiplicadores $L_k(x)$ dependen tan sólo de los puntos $x_0, x_1, \ldots, x_n$ y no de los valores prefijados antes citados. Se llaman *coeficientes de interpolación de Lagrange.* Si utilizamos las fórmulas (15.6) y (15.7) podemos escribir la ecuación (15.9) en la forma

$$(15.10) \qquad L_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j}.$$

Esta fórmula proporciona un método eficiente para calcular el número $L_k(x)$ para un x dado.

Observación. Los coeficientes de Lagrange $L_k(x)$ se expresan frecuentemente en la forma

$$L_k(x) = \frac{A_k(x)}{A'(x_k)},$$

donde A' es la derivada del polinomio (15.5). Para demostrar esa fórmula basta probar que $A'(x_k) = A_k(x_k)$. Derivando la relación

$$A(x) = (x - x_k)A_k(x)$$

obtenemos $A'(x) = (x - x_k)A'_k(x) + A_k(x)$. Cuando $x = x_k$ se obtiene $A'(x_k) = A_k(x_k)$.

EJEMPLO. Determinar el polinomio de grado ≤ 3 que toma los valores y_0, y_1, y_2, y_3 en los puntos -2, -1, 1, 2 respectivamente.

Solución. Tomamos $x_0 = -2$, $x_1 = -1$, $x_2 = 1$, $x_3 = 2$. Los polinomios $L_k(x)$ dados por (15.10) son ahora

$$L_0(x) = \frac{(x+1)(x-1)(x-2)}{(-2+1)(-2-1)(-2-2)} = -\frac{1}{12}(x+1)(x-1)(x-2),$$

$$L_1(x) = \frac{(x+2)(x-1)(x-2)}{(-1+2)(-1-1)(-1-2)} = \frac{1}{6}(x+2)(x-1)(x-2),$$

$$L_2(x) = \frac{(x+2)(x+1)(x-2)}{(1+2)(1+1)(1-2)} = -\frac{1}{6}(x+2)(x+1)(x-2),$$

$$L_3(x) = \frac{(x+2)(x+1)(x-1)}{(2+2)(2+1)(2-1)} = \frac{1}{12}(x+2)(x+1)(x-1).$$

Por lo tanto, el polinomio que se busca es

$$\begin{aligned} P(x) &= y_0L_0(x) + y_1L_1(x) + y_2L_2(x) + y_3L_3(x) \\ &= -\frac{y_0}{12}(x+1)(x-1)(x-2) + \frac{y_1}{6}(x+2)(x-1)(x-2) \\ &\quad - \frac{y_2}{6}(x+2)(x+1)(x-2) + \frac{y_3}{12}(x+2)(x+1)(x-1). \end{aligned}$$

Para calcular el valor de $P(x)$ para un x determinado es mejor dejar el polinomio en esta forma en lugar de ordenarlo en potencias crecientes de x. Por ejemplo, si $y_0 = -5$, $y_1 = 1$, $y_2 = 1$, e $y_3 = 7$, el valor de $P(x)$ para $x = 3/2$ viene dado por

$$\begin{aligned} P(\tfrac{3}{2}) &= \tfrac{5}{12}(\tfrac{5}{2})(\tfrac{1}{2})(-\tfrac{1}{2}) + \tfrac{1}{6}(\tfrac{7}{2})(\tfrac{1}{2})(-\tfrac{1}{2}) - \tfrac{1}{6}(\tfrac{7}{2})(\tfrac{5}{2})(-\tfrac{1}{2}) + \tfrac{7}{12}(\tfrac{7}{2})(\tfrac{5}{2})(\tfrac{1}{2}) \\ &= -\tfrac{25}{96} - \tfrac{7}{48} + \tfrac{35}{48} + \tfrac{245}{96} = \tfrac{276}{96} = 2\tfrac{7}{8}. \end{aligned}$$

15.7 Puntos de interpolación igualmente separados

En la discusión anterior se ha supuesto que los puntos de interpolación $x_0, x_1, \ldots, x_n$ eran distintos, pero por lo demás cualesquiera. Si ahora suponemos que están igualmente separados, vamos a demostrar que los coeficientes de Lagrange se simplifican considerablemente. Supongamos $x_0 < x_1 < x_2 < \ldots < x_n$, y que

h es la distancia entre cada dos puntos consecutivos. Podemos entonces escribir

$$x_j = x_0 + jh$$

para $j = 0, 1, 2, \ldots, n$. Ya que $x_k - x_j = (k - j)h$, la ecuación (15.10) se transforma en

$$L_k(x) = \prod_{\substack{j=0 \\ j\neq k}}^{n} \frac{x - x_0 - jh}{(k - j)h} = \prod_{\substack{j=0 \\ j\neq k}}^{n} \frac{t - j}{k - j}, \tag{15.11}$$

donde

$$t = \frac{x - x_0}{h}.$$

En el último término del segundo miembro de (15.11) el producto de los factores independientes de t es

$$\prod_{\substack{j=0 \\ j\neq k}}^{n} \frac{1}{k - j} = \left(\prod_{j=0}^{k-1} \frac{1}{k - j}\right)\left(\prod_{j=k+1}^{n} \frac{1}{k - j}\right) = \frac{1}{k!} \prod_{j=k+1}^{n} \frac{(-1)}{j - k}$$
$$= \frac{(-1)^{n-k}}{k!\,(n - k)!} = \frac{(-1)^{n-k}}{n!}\binom{n}{k}, \tag{15.12}$$

siendo $\binom{n}{k}$ el coeficiente binomial. Puesto que $x = x_0 + th$, la ecuación (15.11) se escribirá ahora

$$L_k(x_0 + th) = \frac{(-1)^{n-k}}{n!}\binom{n}{k} \prod_{\substack{j=0 \\ j\neq k}}^{n} (t - j). \tag{15.13}$$

Para cada n fijo, el segundo miembro de (15.13) es una función de k y de t que puede tabularse. En el National Bureau of Standards se prepararon tablas de coeficientes de Lagrange para puntos de interpolación igualmente separados. (Véase referencia 13 en la bibliografía del final de este capítulo.) Si x y h se eligen de manera que el número $t = (x - x_0)/h$ es tal que para él han sido tabulados los coeficientes de Lagrange $L_k(x_0 + th)$, el cálculo de $P(x_0 + th)$ se reduce a una multiplicación de $f(x_k)$ por el valor tabulado $L_k(x_0 + th)$, seguida de adición.

15.8 Análisis del error de la interpolación por polinomios

Sea f una función definida en un intervalo $[a, b]$ que contiene los $n + 1$ puntos $x_0, x_1, \ldots, x_n$, y sea P el polinomio de interpolación de grado $\leq n$ que coincide con f en aquellos puntos. Si alteramos los valores de f en puntos distintos de

los de interpolación no variará el polinomio P. Esto prueba que la función f y el polinomio P pueden diferir considerablemente en los puntos distintos a los de interpolación. Si la función dada f tiene ciertas cualidades de «regularidad» en el intervalo $[a, b]$ podemos esperar que el polinomio de interpolación P será una buena aproximación de f en los puntos distintos de los x_k. El teorema que sigue nos da una útil expresión que nos permite estudiar el error en la interpolación por polinomios cuando la función dada posee derivada de orden $n + 1$ en $[a, b]$.

TEOREMA 15.3. *Sean $x_0, x_1, \ldots, x_n$, $n + 1$ puntos distintos en el dominio de una función f, y P el polinomio de interpolación de grado $\leq n$ que coincide con f en esos puntos. Elijamos un punto x en el dominio de f y sea $[\alpha, \beta]$ un intervalo cerrado cualquiera que contenga los puntos $x_0, x_1, \ldots, x_n$, y x. Si f posee derivada de orden $n + 1$ en el intervalo $[\alpha, \beta]$ existe por lo menos un punto c en el intervalo abierto (α, β) tal que*

$$(15.14) \qquad f(x) - P(x) = \frac{A(x)}{(n+1)!} f^{(n+1)}(c),$$

donde

$$A(x) = (x - x_0)(x - x_1) \cdots (x - x_n).$$

Observación. El punto c depende a la vez de x y de n.

Demostración. Si x es uno de los puntos de interpolación x_k, entonces $A(x_k) = 0$ y la ecuación (15.14) se satisface para cualquier c en (α, β). Supongamos, pues, que x no es ninguno de los puntos de interpolación. Mantengamos fijo x y definamos una nueva función F en $[\alpha, \beta]$ mediante la ecuación

$$(15.15) \qquad F(t) = A(x)[f(t) - P(t)] - A(t)[f(x) - P(x)].$$

El segundo miembro de la ecuación, como función de t, posee derivada de orden $n + 1$; luego lo mismo le ocurre al primer miembro. Puesto que $P(t)$ es un polinomio en t de grado $\leq n$, su derivada $(n + 1)$-ésima es idénticamente nula. El polinomio $A(t)$ es de grado $n + 1$, siendo el término de mayor grado t^{n+1}, y tenemos $A^{(n+1)}(t) = (n + 1)!$. Por lo tanto, si derivamos la ecuación (15.15) $n + 1$ veces respecto a t obtenemos la fórmula

$$(15.16) \qquad F^{(n+1)}(t) = A(x)f^{(n+1)}(t) - (n + 1)!\,[f(x) - P(x)].$$

De la definición que se da en la ecuación (15.15) vemos que F tiene el valor cero en los $n + 1$ puntos de interpolación $x_0, x_1, \ldots, x_n$, y *también* en el punto x. Por consiguiente, $F(t) = 0$ en $n + 2$ puntos distintos del intervalo $[\alpha, \beta]$. Esos

puntos determinan $n + 1$ subintervalos consecutivos de $[\alpha, \beta]$ y la función F se anula en los dos extremos de esos subintervalos. Según el teorema de Rolle, la derivada $F'(t)$ debe anularse por lo menos en un punto t interior a cada subintervalo. Si elegimos exactamente uno de tales puntos t en cada subintervalo obtenemos $n + 1$ puntos distintos en el intervalo abierto (α, β) en los que $F'(t) = 0$. Esos puntos, a su vez, determinan n subintervalos en cuyos extremos tenemos $F'(t) = 0$. Aplicando el teorema de Rolle a F' encontramos que la derivada segunda $F''(t)$ es cero por lo menos en n puntos distintos de (α, β). Aplicando reiteradamente el teorema de Rolle $n + 1$ veces de ese modo encontramos finalmente que existe por lo menos un punto c en (α, β) en el que $F^{(n+1)}(c) = 0$. Sustituyendo este valor de c en la ecuación (15.16) obtenemos

$$(n + 1)!\,[f(x) - P(x)] = A(x)f^{(n+1)}(c),$$

que coincide con (15.14). Con lo que queda demostrado el teorema.

Debe observarse que, al igual que en la aproximación con polinomios de Taylor, el término que da el error contiene la derivada $(n + 1)$-ésima $f^{(n+1)}(c)$ calculada en un punto desconocido c. Si se conocen los valores extremos de $f^{(n+1)}$ en $[\alpha, \beta]$, se pueden obtener cotas superior e inferior del error.

Supongamos ahora que los puntos de interpolación están igualmente separados y que $x_0 < x_1 < x_2 < \ldots < x_n$. Si h representa la distancia entre cada dos consecutivos podemos escribir

$$x_j = x_0 + jh \qquad \text{y} \qquad x = x_0 + th\,,$$

donde $t = (x - x_0)/h$. Ya que $x - x_j = (t - j)h$, el polinomio $A(x)$ puede escribirse en la forma siguiente

$$A(x) = \prod_{j=0}^{n} (x - x_j) = h^{n+1} \prod_{j=0}^{n} (t - j).$$

La fórmula (15.14) se transforma ahora en

$$\text{(15.17)} \qquad f(x) - P(x) = \frac{f^{(n+1)}(c)}{(n + 1)!}\, h^{n+1} \prod_{j=0}^{n} (t - j),$$

con $t = (x - x_0)/h$.

EJEMPLO. *El error en la interpolación lineal.* Supongamos que una función f con derivada segunda está tabulada y que deseamos estimar su valor en un punto x intermedio entre los dos consecutivos x_0 y $x_0 + h$. Si hacemos uso de la interpolación lineal aproximamos la gráfica de f en el intervalo $[x_0, x_0 + h]$ por

una recta, como muestra la figura 15.1. Si P es el polinomio de interpolación de primer grado, el error estimado en (15.17) es ahora

$$f(x) - P(x) = \frac{f''(c)}{2!} h^2 t(t-1), \tag{15.18}$$

donde $t = (x - x_0)/h$. Cuando x está situado entre x_0 y $x_0 + h$ tenemos $0<t<1$ y el valor máximo de $|t(t-1)|$ en este intervalo es $\frac{1}{4}$. Por consiguiente (15.18) nos da la estimación

$$|f(x) - P(x)| \leq \frac{|f''(c)|\, h^2}{8}.$$

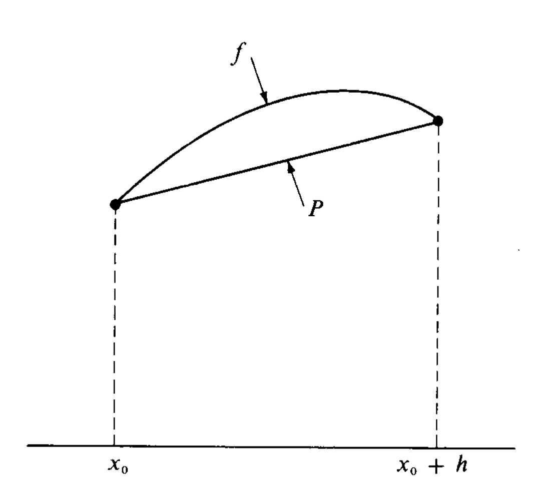


FIGURA 15.1 *Interpolación lineal.*

El punto c es un punto desconocido en el intervalo $(x_0, x_0 + h)$. Si la derivada segunda f'' está acotada en ese intervalo, $|f''(x)| \leq M$, la estimación del error es

$$|f(x) - P(x)| \leq \frac{Mh^2}{8}.$$

En particular, si f es un seno o un coseno; entonces $|f''(x)| \leq 1$ para todo x y tenemos $|f(x)-P(x)| \leq h^2/8$. Con una tabla de senos o cosenos con intervalos de grado en grado (un grado$=\pi/180$ radianes) tenemos $h=\pi/180$, así que

$$\frac{h^2}{8} = \frac{\pi^2}{8(180)^2} < \frac{10}{259{,}200} < \frac{1}{25{,}000} = 0.00004.$$

Ya que este error no excede a ½ en la cuarta cifra decimal, la interpolación lineal en una tabla de cuatro cifras decimales sería satisfactoria. El error estimado puede mejorarse en las partes de la tabla en donde $|f''(c)|$ sea menor que 1.

15.9 Ejercicios

1. En cada caso encontrar el polinomio P de menor grado que satisfaga las condiciones dadas.

   a) $P(-1) = 0, \quad P(0) = 2, \quad P(2) = 7.$
   b) $P(1) = 1, \quad P(2) = 0, \quad P(3) = 0, \quad P(4) = 1.$
   c) $P(1) = 1, \quad P(2) = 2, \quad P(3) = 3, \quad P(0) = 1.$
   d) $P(0) = -2, \quad P(1) = 0, \quad P(-1) = -2, \quad P(2) = 16.$
   e) $P(-2) = 11, \quad P(-1) = -11, \quad P(0) = -5, \quad P(1) = -1.$

2. Sea $f(x) = \cos(\pi x/4)$. Hallar el polinomio de menor grado que toma los mismos valores que f en los puntos $-2, -\frac{4}{3}, 0, \frac{4}{3}, 2$.
3. Sean P un polinomio de grado $\leq n$ y $A(x) = (x - x_0)(x - x_1) \ldots (x - x_n)$, siendo $x_0, x_1, \ldots, x_n$, $n + 1$ puntos distintos.
   a) Demostrar que cualquiera que sea el polinomio B el polinomio Q dado por $Q(x) = P(x) + A(x)B(x)$ coincide con P en los puntos $x_0, x_1, \ldots, x_n$.
   b) Probar también el recíproco. Esto es, si Q es un polinomio cualquiera que coincide con P en los puntos $x_0, x_1, \ldots x_n$, $Q(x) = P(x) + A(x)B(x)$ para un cierto polinomio B.
4. a) Hallar el polinomio Q de menor grado que satisfaga las condiciones

$$Q(-2) = -5, \qquad Q(-1) = -1, \qquad Q(1) = 1, \qquad Q'(0) = -1.$$

   [*Indicación:* Hallar primero el polinomio P que tome los valores prescritos en $-2, -1, 1$, y utilizar luego el ejercicio 3 para determinar Q.]
   b) Hallar el polinomio Q de menor grado que satisfaga las condiciones de la parte a) con $Q'(0) = -3$ en lugar de $Q'(0) = -1$.
5. Sea $f(x) = \log_4 x$ para $x > 0$. Calcular $P(32)$, siendo P el polinomio de menor grado que coincide con f en los puntos:

   a) $x = 1, 64.$
   b) $x = 1, 16, 256.$
   c) $x = 4, 16, 64.$
   d) $x = 1, 4, 16, 64, 256.$

   Calcular en cada caso la diferencia $f(32) - P(32)$. Estos ejemplos prueban que la precisión en la interpolación por polinomios no se mejora necesariamente aumentando el número de puntos de interpolación.
6. Los coeficientes de interpolación de Lagrange $L_k(x)$ dados por la ecuación (15.10) dependen no solamente de x, sino también de los puntos de interpolación $x_0, x_1, \ldots, x_n$. Esta dependencia podemos indicarla escribiendo $L_k(x) = L_k(x; \boldsymbol{X})$, donde $\boldsymbol{X}$ representa el vector del espacio E_{n+1} dado por $\boldsymbol{X} = (x_0, x_1, \ldots, x_n)$. Para un número real dado b, representemos con $\boldsymbol{b}$ el vector de E_{n+1} cuyos componentes son todos iguales a b. Si $a \neq 0$, demostrar que

$$L_k(ax + b\,;\; a\boldsymbol{X} + \boldsymbol{b}) = L_k(x; \boldsymbol{X}).$$

   Ésta es la llamada *propiedad de invariancia* de los coeficientes de interpolación de La-

grange. El ejercicio siguiente nos hace ver la utilidad de esta propiedad para simplificar los cálculos.

7. Sea P el polinomio de grado ≤ 4 que toma los valores

$$P(2.4) = 72, \quad P(2.5) = 30, \quad P(2.7) = 18, \quad P(2.8) = 24, \quad P(3.0) = 180.$$

a) Introducir nuevos puntos de interpolación u_j ligados a los dados x_j por la ecuación $u_j = 10x_j - 24$. Los u_j son enteros. Para cada $k = 0, 1, 2, 3, 4$, determinar los coeficientes de interpolación de Lagrange $L_k(x)$ en función de u, siendo $u = 10x - 24$.
b) Utilizar la propiedad de invariancia del ejercicio 6 para calcular $P(2,6)$.

8. Una tabla de la función $f(x) = \log x$ contiene entradas desde $x = 1$ hasta $x = 10$ a intervalos de 0,001. Los valores intermedios a cada par de entradas consecutivas se han calculado mediante interpolación lineal. Supongamos que las entradas en la tabla son exactas.
a) Demostrar que el error en la interpolación lineal no excederá a $\frac{1}{8}$ en la sexta cifra decimal.
b) ¿Para qué valores de x será satisfactoria la interpolación lineal para una tabla de siete cifras?
c) ¿Cuál debería ser la separación de las entradas en el intervalo $1 \leq x \leq 2$ de modo que la interpolación lineal sea satisfactoria en una tabla de siete cifras?

En los ejercicios del 9 al 15, $x_0, x_1, \ldots, x_n$ son puntos distintos y

$$A(x) = \prod_{j=0}^{n} (x - x_j), \qquad A_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} (x - x_j), \qquad L_k(x) = \frac{A_k(x)}{A_k(x_k)}.$$

9. Deducir la fórmula $A'(x) = \sum_{k=0}^{n} A_k(x)$ mediante el uso, a) de la derivación logarítmica, b) de la fórmula de interpolación de Lagrange.

10. Demostrar cada una de las fórmulas siguientes:

a) $\displaystyle\sum_{k=0}^{n} L_k(x) = 1$ y $\displaystyle\sum_{k=0}^{n} \frac{A_k'(x)}{A'(x_k)} = 0$ para todo x.

b) $\displaystyle\sum_{k=0}^{n} \frac{1}{A'(x_k)} = 0$. [*Indicación:* Utilizar la parte a) con valores de x convenientes.]

11. Sea P un polinomio cualquiera de grado $\leq n$. Demostrar que el coeficiente de x^n es igual a

$$\sum_{k=0}^{n} \frac{P(x_k)}{A'(x_k)}.$$

12. a) Determinar a y b de modo que el polinomio

$$P_k(x) = \{a + b(x - x_k)\}L_k(x)^2$$

tenga las siguientes propiedades:

$$P_k(x_i) = 0 \text{ para todo } i, \qquad P_k'(x_k) = 1, \qquad \text{y} \qquad P_k'(x_i) = 0 \quad \text{para} \quad i \neq k.$$

b) Determinar c y d de modo que el polinomio

$$Q_k(x) = \{c + d(x - x_k)\}L_k(x)^2$$

tenga las siguientes propiedades:

$$Q_k(x_k) = 1, \quad Q_k(x_i) = 0 \quad \text{para} \quad i \neq k, \quad \text{y} \quad Q_k'(x_i) = 0 \text{ para todo } i.$$

c) Sea $H(x) = \sum_{k=0}^{n} f(x_k)Q_k(x) + \sum_{k=0}^{n} f'(x_k)P_k(x)$ donde f es una función dada derivable en $x_0, x_1, \ldots, x_n$. Demostrar que

$$H(x_i) = f(x_i) \quad \text{y} \quad H'(x_i) = f'(x_i) \text{ para todo } i.$$

Demostrar también que existe a lo más un polinomio $H(x)$ de grado $\leq 2n + 1$ con esa propiedad.

13. Dados $n + 1$ puntos distintos $x_0, x_1, \ldots, x_n$, sean P y Q dos polinomios de grado $\leq n$ que satisfagan las $n + 1$ condiciones

$$P(x_0) = Q(x_0), \quad P'(x_1) = Q'(x_1), \quad P''(x_2) = Q''(x_2), \quad \ldots, \quad P^{(n)}(x_n) = Q^{(n)}(x_n).$$

Demostrar que $P(x) = Q(x)$ para todo x.
b) Sea $B_0(x) = 1$, y para $n \geq 1$ definimos

$$B_n(x) = \frac{x(x - n)^{n-1}}{n!}.$$

Demostrar que $B_n'(x) = B_{n-1}(x - 1)$ para $n \geq 1$ y deducir que

$$B_n(0) = B_n'(1) = B_n''(2) = \cdots = B_n^{(n-1)}(n - 1) = 0 \quad \text{y} \quad B_n^{(n)}(n) = 1.$$

c) Demostrar que el único polinomio de grado $\leq n$ que satisface las condiciones

$$P(0) = c_0, \quad P'(1) = c_1, \quad P''(2) = c_2, \quad \ldots, \quad P^{(n)}(n) = c_n$$

viene dado por

$$P(x) = \sum_{k=0}^{n} c_k B_k(x).$$

d) Si $x_k = x_0 + kh$ para $k = 0, 1, 2, \ldots, n$, donde $h > 0$, generalizar los resultados de b) y c).

14. Suponiendo que $x_0, x_1, \ldots, x_n$ son enteros que satisfacen $x_0 < x_1 < \ldots < x_n$.
a) Demostrar que $|A'(x_k)| \geq k!\,(n - k)!$ y deducir que

$$\sum_{k=0}^{n} \frac{1}{|A'(x_k)|} \leq \frac{2^n}{n!}.$$

b) Sea P un polinomio cualquiera de grado n, cuyo término de mayor grado es igual a x^n. Sea M el mayor de los números $|P(x_0)|$, $|P(x_1)|$, ..., $|P(x_n)|$. Demostrar que $M \geqslant n!/2^n$. *[Indicación:* Utilizar la parte a) del ejercicio 11.]

15. Demostrar las siguientes fórmulas. En las partes a) y b), x es un punto cualquiera distinto de $x_0, x_1, \ldots, x_n$.

a) $$\frac{A'(x)}{A(x)} = \sum_{j=0}^{n} \frac{1}{x - x_j}.$$

b) $$\frac{A''(x)}{A'(x)} = \frac{A_k(x)}{A'(x)} \sum_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{x - x_j} + \sum_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{x - x_j} - \frac{A(x)}{A'(x)} \sum_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{(x - x_j)^2}.$$

c) $$\frac{A''(x_k)}{A'(x_k)} = 2 \sum_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{x_k - x_j}.$$

16. Sea $P_n(x)$ el polinomio de grado $\leq n$ que coincida con la función $f(x) = e^{ax}$ en los $n + 1$ enteros $x = 0, 1, \ldots, n$. Puesto que ese polinomio depende de a lo designaremos por $P_n(x; a)$. Demostrar que el límite

$$\lim_{a \to 0} \frac{P_n(x;\ a) - 1}{a}$$

existe y es un polinomio en x. Determinar explícitamente dicho polinomio.

15.10 Fórmula de interpolación de Newton

Representemos con P_n el polinomio de interpolación de grado $\leq n$ que coincide con una función dada f en $n + 1$ puntos distintos $x_0, x_1, \ldots, x_n$. La fórmula de interpolación de Lagrange nos dice que

$$P_n(x) = \sum_{k=0}^{n} L_k(x) f(x_k),$$

donde $L_k(x)$ es un polinomio de grado n (el coeficiente de interpolación de Lagrange) dado por el producto

(15.19) $$L_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j}, \quad \text{para} \quad k = 0, 1, 2, \ldots, n.$$

Supongamos que añadimos un nuevo punto de interpolación x_{n+1} a los dados $x_0, x_1, \ldots, x_n$. Para determinar el polinomio correspondiente P_{n+1} con la fórmula de Lagrange es necesario calcular un nuevo coeficiente de interpolación L_{n+1} y

recalcular todos los coeficientes anteriores $L_0, L_1, \ldots, L_n$, cada uno de los cuales es ahora un polinomio de grado $n + 1$. En la práctica esto supone un trabajo considerable. Por consiguiente, es conveniente disponer de otra fórmula para determinar P_n que permita una más fácil transición de P_n a P_{n+1}. Tal fórmula la descubrió Newton; la deduciremos a partir del teorema siguiente.

TEOREMA 15.4. *Dados $n + 2$ puntos distintos $x_0, x_1, \ldots, x_n, x_{n+1}$. Sea P_n el polinomio de grado $\leq n$ que coincide con una función dada f en $x_0, \ldots, x_n$, y sea P_{n+1} el polinomio de grado $\leq n + 1$ que coincide con f en $x_0, x_1, \ldots, x_n, x_{n+1}$. Existe entonces una constante c_{n+1}, determinada con unicidad por f y por los puntos de interpolación $x_0, \ldots, x_{n+1}$ tal que*

$$P_{n+1}(x) = P_n(x) + c_{n+1}(x - x_0) \cdots (x - x_n). \tag{15.20}$$

Demostración. Pongamos $Q(x) = P_n(x) + c(x - x_0) \ldots (x - x_n)$, donde c es una constante no determinada. Entonces Q es un polinomio de grado $\leq n + 1$ que coincide con P_n y por tanto con f en cada uno de los $n + 1$ puntos $x_0, \ldots, x_n$. Elijamos luego c de modo que Q coincida también con f en x_{n+1}. Esto exige

$$f(x_{n+1}) = P_n(x_{n+1}) + c(x_{n+1} - x_0) \cdots (x_{n+1} - x_n).$$

Puesto que el coeficiente c no es nulo, esta ecuación tiene una sola solución que llamamos c_{n+1}. Tomando $c = c_{n+1}$ vemos que $Q = P_{n+1}$.

El teorema que sigue expresa $P_n(x)$ en función de los números $c_1, \ldots, c_n$.

TEOREMA 15.5. FÓRMULA DE INTERPOLACIÓN DE NEWTON. *Si $x_0, \ldots, x_n$ son distintos, tenemos*

$$P_n(x) = f(x_0) + \sum_{k=1}^{n} c_k(x - x_0) \cdots (x - x_{k-1}). \tag{15.21}$$

Demostración. Definamos $P_0(x) = f(x_0)$ y tomemos $n = 0$ en (15.20) obteniendo

$$P_1(x) = f(x_0) + c_1(x - x_0).$$

Hagamos ahora $n = 1$ en (15.20), con lo que conseguimos

$$P_2(x) = P_1(x) + c_2(x - x_0)(x - x_1) = f(x_0) + c_1(x - x_0) + c_2(x - x_0)(x - x_1).$$

Por inducción, obtenemos (15.21).

La propiedad de la fórmula de Newton expresada en la ecuación (15.20)

nos permite calcular P_{n+1} con sólo sumar un nuevo término a P_n. Esta propiedad no la posee la fórmula de Lagrange.

La utilidad de la fórmula de Newton depende, naturalmente, de la facilidad con la que puedan calcularse los coeficientes $c_1, c_2, \ldots, c_n$. El teorema que sigue demuestra que c_n es una combinación lineal de los valores $f(x_0), \ldots, f(x_n)$.

TEOREMA 15.6. *Los coeficientes de la fórmula de interpolación de Newton vienen dados por*

$$(15.22) \qquad c_n = \sum_{k=0}^{n} \frac{f(x_k)}{A_k(x_k)}, \qquad \text{donde} \quad A_k(x_k) = \prod_{\substack{j=0 \\ j \neq k}}^{n} (x_k - x_j).$$

Demostración. Según la fórmula de Lagrange tenemos

$$P_n(x) = \sum_{k=0}^{n} L_k(x) f(x_k),$$

en donde $L_k(x)$ es el polinomio de grado n dado por (15.19). Puesto que el coeficiente de x^n en $L_k(x)$ es $1/A_k(x_k)$, el coeficiente de x^n en $P_n(x)$ es la suma que aparece en (15.22). Por otra parte, la fórmula de Newton demuestra que el coeficiente de x^n en $P_n(x)$ es igual a c_n. Esto completa la demostración.

La ecuación (15.22) nos proporciona un método directo para calcular los coeficientes de la fórmula de Newton. Los números $A_k(x_k)$ aparecen también como factores en el denominador del coeficiente de interpolación de Lagrange $L_k(x)$. La sección que sigue describe otro método para calcular los coeficientes cuando los puntos de interpolación están igualmente separados.

15.11 Puntos de interpolación igualmente separados. El operador de las diferencias sucesivas

En el caso de puntos de interpolación igualmente separados $x_k = x_0 + kh$ para $k = 0, 1, \ldots, n$ podemos utilizar la ecuación (15.12) obteniendo

$$\frac{1}{A_k(x_k)} = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{x_k - x_j} = \frac{1}{h^n} \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{k - j} = \frac{(-1)^{n-k}}{n!\, h^n} \binom{n}{k}.$$

En este caso la fórmula que da c_n del teorema 15.6 se transforma en

$$(15.23) \qquad c_n = \frac{1}{n!\, h^n} \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} f(x_k),$$

La suma del segundo miembro puede calcularse de otra manera en función de un operador lineal Δ llamado *operador de las diferencias sucesivas.*

DEFINICIÓN. *Sean h un número real fijo y f una función dada. La función Δf definida por la ecuación*

$$\Delta f(x) = f(x + h) - f(x)$$

se llama la diferencia primera de f. Está definida para aquellos puntos x para los que tanto x como $x + h$ están en el dominio de f. Las diferencias de órdenes superiores $\Delta^2 f$, $\Delta^3 f$, ... se definen por inducción como siguen:

$$\Delta^{k+1} f = \Delta(\Delta^k f) \quad \textit{para} \quad k = 1, 2, 3, \ldots .$$

Observación. También se emplean las notaciones $\Delta_h f(x)$ y $\Delta f(x; h)$ cuando se quiere indicar la dependencia de h. Es conveniente definir $\Delta^0 f = f$.

La n-ésima diferencia $\Delta^n f(x)$ es una combinación lineal de los valores $f(x)$, $f(x + h)$, ..., $f(x + nh)$. Por ejemplo, tenemos

$$\begin{aligned} \Delta^2 f(x) &= \{f(x + 2h) - f(x + h)\} - \{f(x + h) - f(x)\} \\ &= f(x + 2h) - 2f(x + h) + f(x). \end{aligned}$$

En general, tenemos

$$\Delta^n f(x) = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} f(x + kh). \tag{15.24}$$

Esto se demuestra fácilmente por inducción respecto a n, utilizando la ley del triángulo de Pascal para los coeficientes binomiales:

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}.$$

Supongamos ahora que f esté definida en $n + 1$ puntos igualmente separados $x_k = x_0 + kh$ para $k = 0, 1, \ldots, n$. Entonces de (15.23) y (15.24) obtenemos la fórmula

$$c_n = \frac{1}{h^n n!} \Delta^n f(x_0).$$

Esto proporciona un método rápido para calcular los coeficientes de la fórmula de interpolación de Newton. El diagrama de la tabla 15.1, llamado *tabla de diferencias,* hace ver cómo las diferencias sucesivas pueden calcularse sistemáti-

camente a partir de una tabulación de los valores de f en los puntos igualmente separados. En la tabla se ha escrito f_k en lugar de $f(x_k)$.

La fórmula de interpolación de Newton (15.21) toma ahora la forma

$$P_n(x) = f(x_0) + \sum_{k=1}^{n} \frac{\Delta^k f(x_0)}{k!\ h^k} \prod_{j=0}^{k-1} (x - x_j). \tag{15.25}$$

TABLA 15.1

x	$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$
x_0	f_0			
		$f_1 - f_0 = \Delta f(x_0)$		
x_1	f_1		$\Delta f(x_1) - \Delta f(x_0) = \Delta^2 f(x_0)$	
		$f_2 - f_1 = \Delta f(x_1)$		$\Delta^2 f(x_1) - \Delta^2 f(x_0) = \Delta^3 f(x_0)$
x_2	f_2		$\Delta f(x_2) - \Delta f(x_1) = \Delta^2 f(x_1)$	
		$f_3 - f_2 = \Delta f(x_2)$		
x_3	f_3			

Si escribimos

$$\prod_{j=0}^{k-1} (x - x_j) = \prod_{j=0}^{k-1} (x - x_0 - jh) = h^k \prod_{j=0}^{k-1} \left(\frac{x - x_0}{h} - j\right) = h^k \prod_{j=0}^{k-1} (t - j),$$

en donde $t = (x - x_0)/h$, la ecuación (15.25) se convierte en

$$P_n(x) = f(x_0) + \sum_{k=1}^{n} \frac{\Delta^k f(x_0)}{k!} \prod_{j=0}^{k-1} (t - j). \tag{15.26}$$

15.12 Polinomios factoriales

Si k es un número natural, el producto $t(t - 1) \ldots (t - k + 1)$ es un poli-

nomio en t de grado k llamado *polinomio factorial* k-ésimo. Se representa con el símbolo $t^{(k)}$. Así pues, según la definición,

$$t^{(k)} = \prod_{j=0}^{k-1} (t - j).$$

Definimos también $t^{(0)} = 1$. Si consideramos el operador Δ de las diferencias sucesivas con $h = 1$, es decir, $\Delta f(x) = f(x + 1) - f(x)$, encontramos que

$$\Delta t^{(n)} = n t^{(n-1)} \quad \text{para} \quad n \geq 1 .$$

Es una fórmula del mismo tipo que la de la derivación ordinaria de potencias $Dt^n = nt^{n-1}$. Con lo cual vemos que el polinomio factorial $t^{(n)}$ está relacionado con las diferencias del mismo modo que las potencias ordinarias t^n lo están con las derivadas.

Con el uso de los polinomios factoriales, la fórmula de interpolación de Newton (8.65) se convierte en

$$P_n(x_0 + th) = \sum_{k=0}^{n} \frac{\Delta^k f(x_0)}{k!} t^{(k)} .$$

Puesta en esta forma, la fórmula de Newton se parece a la de Taylor para el polinomio de grado $\leq n$ que coincide con f y sus n primeras derivadas en x_0. Si escribimos

$$\binom{t}{k} = \frac{t^{(k)}}{k!} = \frac{t(t-1)\cdots(t-k+1)}{k!},$$

la fórmula de Newton toma la forma

$$P_n(x_0 + th) = \sum_{k=0}^{n} \binom{t}{k} \Delta^k f(x_0) .$$

En los ejercicios que siguen se exponen otras propiedades de los polinomios factoriales.

15.13 Ejercicios

1. Sea $\Delta f(x) = f(x + h) - f(x)$. Si f es un polinomio de grado n,

$$f(x) = \sum_{r=0}^{n} a_r x^r$$

con $a_n \neq 0$, demostrar que *a*) $\Delta^k f(x)$ es un polinomio de grado $n-k$ si $k \le n$; b) $\Delta^n f(x) = n!h^n a_n$; c) $\Delta^k f(x) = 0$ para $k > n$.

2. Sea $\Delta f(x) = f(x+h) - f(x)$. Si $f(x) = \operatorname{sen}(ax+b)$, demostrar que

$$\Delta^n f(x) = \left(2 \operatorname{sen} \frac{ah}{2}\right)^n \operatorname{sen}\left(ax + b + \frac{nah + n\pi}{2}\right).$$

3. Sea $\Delta f(x) = f(x+h) - f(x)$.
a) Si $f(x) = a^x$, siendo $a > 0$, demostrar que $\Delta^k f(x) = (a^h - 1)^k a^x$.
b) Si $g(x) = (1+a)^{x/h}$, siendo $a > 0$ demostrar que $\Delta^k g(x) = a^k g(x)$.
c) Demostrar que el polinomio P_n de grado n que toma los valores $P_n(k) = (1+a)^k$ para $k = 0, 1, 2, \ldots, n$ viene dado por

$$P_n(x) = \sum_{k=0}^{n} \frac{a^k}{k!} x^{(k)}.$$

4. Sea $x^{(n)}$ el n-ésimo polinomio factorial en x. Puesto que $x^{(n)}$ es un polinomio en x de grado n con el valor 0 cuando $x = 0$, podemos escribir

$$x^{(n)} = \sum_{k=1}^{n} S_{k,n} x^k.$$

Los números $S_{k,n}$ se llaman ***números de Stirling de primera especie.*** De la definición de $x^{(n)}$ resulta evidente que $S_{n,n} = 1$ para $n \geq 0$.
a Demostrar que $S_{n-1,n} = -n(n-1)/2$ y que $S_{1,n} = (-1)^{n-1}(n-1)!$ para $n \geq 1$.
b) Demostrar que $S_{k,n+1} = S_{k-1,n} - nS_{k,n}$. Comprobar esta relación en la tabla 15.2, de números de Stirling de primera especie, y construir las tres filas siguientes de la tabla.

TABLA 15.2

n	$S_{1,n}$	$S_{2,n}$	$S_{3,n}$	$S_{4,n}$	$S_{5,n}$	$S_{6,n}$	$S_{7,n}$
1	1						
2	−1	1					
3	2	−3	1				
4	−6	11	−6	1			
5	24	−50	35	−10	1		
6	−120	274	−225	85	−15	1	
7	720	−1764	1624	−735	175	−21	1

c) Expresar el polinomio $x^{(4)} + 3x^{(3)} + 2x^{(1)} + 1$ como combinación lineal de potencias de x.

5. a) Demostrar que

$$x = x^{(1)}, \qquad x^2 = x^{(1)} + x^{(2)}, \qquad x^3 = x^{(1)} + 3x^{(2)} + x^{(3)},$$

y que, en general,

$$x^n = \sum_{k=1}^{n} \frac{\Delta^k f(0)}{k!} x^{(k)},$$

donde $f(x) = x^n$ y $\Delta f(x) = f(x + 1) - f(x)$. Los números $T_{k,n} = \Delta^k f(0)/k!$ se llaman *números de Stirling de segunda especie.*
b) Demostrar que

$$\Delta^k x^{n+1} = (x + k)\, \Delta^k x^n + k\, \Delta^{k-1} x^n$$

y con ello deducir que $T_{k,n+1} = T_{k-1,n} + kT_{k,n}$.
c) Comprobar la tabla 15.3 de números de Stirling de segunda especie, con la fórmula de recurrencia de la parte b), y construir las tres filas siguientes de la tabla.

TABLA 15.3

n	$T_{1,n}$	$T_{2,n}$	$T_{3,n}$	$T_{4,n}$	$T_{5,n}$	$T_{6,n}$	$T_{7,n}$
1	1						
2	1	1					
3	1	3	1				
4	1	7	6	1			
5	1	15	25	10	1		
6	1	31	90	65	15	1	
7	1	63	301	350	140	21	1

d) Expresar el polinomio $x^4 + 3x^3 + 2x - 1$ como una combinación lineal de polinomios factoriales.

6. a) Si p es un número natural y a y b son enteros tales que $a < b$, demostrar que

$$\sum_{k=a}^{b-1} k^{(p)} = \frac{b^{(p+1)} - a^{(p+1)}}{p + 1}.$$

Esta fórmula es análoga a la de la integración $\int_a^b x^p\, dx$. Debe observarse, sin embargo, que el límite superior de la suma es $b - 1$, y no b.
b) Comprobar que $k(k + 3) = 4k^{(1)} + k^{(2)}$. Usar entonces la parte a) para demostrar que

$$\sum_{k=1}^{n} k(k + 3) = 4\frac{(n + 1)^{(2)}}{2} + \frac{(n + 1)^{(3)}}{3} = \frac{n(n + 1)(n + 5)}{3}.$$

c) Si $f(k)$ es un polinomio en k de grado r, demostrar que

$$\sum_{k=1}^{n} f(k)$$

es un polinomio en n de grado $r + 1$.

7. Con el método sugerido en el ejercicio 6 expresar cada una de las sumas siguientes como un polinomio en n.

a) $\sum_{k=1}^{n} (4k^2 + 7k + 6)$.

b) $\sum_{k=1}^{n} k^2(k + 1)$.

c) $\sum_{k=1}^{n} k(k + 1)(k + 2)$.

d) $\sum_{k=1}^{n} k^4$.

8. Designemos con A el operador lineal definido por la ecuación

$$A(f) = a_0 \Delta^n f + a_1 \Delta^{n-1} f + \cdots + a_{n-1} \Delta f + a_n f,$$

siendo $a_0, a_1, \ldots, a_n$ constantes. Éste es el llamado *operador diferencia de coeficientes constantes*. Es análogo al operador derivada de coeficientes constantes descrito en la sección 6.7. A cada uno de tales A podemos asociar el polinomio característico p_A definido por

$$p_A(r) = a_0 r^n + a_1 r^{n-1} + \cdots + a_{n-1} r + a_n .$$

Recíprocamente, a todo polinomio p podemos asociar un operador A del cual es polinomio característico. Si A y B son dos operadores diferencia de coeficientes constantes y λ es un número real fijo, definimos $A + B$, AB, y λA mediante las mismas fórmulas empleadas en la sección 6.7 para los operadores derivada. Demostrar entonces que el teorema 6.6 es válido para operadores diferencia de coeficientes constantes.

15.14 Problema de mínimo relativo a la norma del máximo

Consideremos un problema que se presenta espontáneamente en la teoría de la aproximación por polinomios. En el teorema 15.3 se dedujo la fórmula del error

$$f(x) - P(x) = \frac{A(x)}{(n + 1)!} f^{(n+1)}(c), \tag{15.27}$$

en donde

$$A(x) = (x - x_0)(x - x_1) \cdots (x - x_n).$$

Aquí P es el único polinomio de grado $\leq n$ que coincide con f en $n + 1$ puntos

distintos $x_0, x_1, \ldots, x_n$ en $[a, b]$. La función f se supone con derivada de orden $n+1$ en $[a, b]$, y c es un punto desconocido situado en $[a, b]$. Para estimar el error en (15.27) son necesarias cotas para la derivada $f^{(n+1)}$ y para el producto $A(x)$. Como A es un polinomio, su valor absoluto posee un máximo en algún punto del intervalo $[a, b]$. Este máximo dependerá de la elección de los puntos $x_0, x_1, \ldots, x_n$, y es natural procurar elegir esos puntos de manera que la norma A sea la menor posible.

Podemos designar ese máximo por $\|A\|$, donde $\|A\|$ es la norma del máximo, dada por

$$\|A\| = \max_{a \leq x \leq b} |A(x)| .$$

El problema consiste en encontrar un polinomio de grado asignado que haga mínimo $\|A\|$. El primero en resolver este problema fue Chebyshev; su solución conduce a una clase interesante de polinomios que también se presentan en otras cuestiones. Primero daremos una breve noticia sobre esos polinomios y luego volveremos al problema de mínimo indicado.

15.15 Polinomios de Chebyshev

Sea $x + iy$ un número complejo de valor absoluto 1. En virtud del desarrollo de Newton tenemos

$$(x + iy)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k}(iy)^k$$

para todo entero $n \geq 0$. Hagamos en esta fórmula $x = \cos\theta$, $y = \operatorname{sen}\theta$, y consideremos la parte real de cada miembro. Puesto que

$$(x + iy)^n = (\cos\theta + i \operatorname{sen}\theta)^n = e^{in\theta} = \cos n\theta + i \operatorname{sen} n\theta ,$$

identificando esas dos partes tenemos

(15.28) $$\cos n\theta = x^n - \binom{n}{2} x^{n-2} y^2 + \binom{n}{4} x^{n-4} y^4 - + \cdots .$$

Ya que $y^2 = \operatorname{sen}^2\theta = 1 - \cos^2\theta = 1 - x^2$, el segundo miembro de (15.28) es un polinomio en x de grado n. Ese polinomio es el polinomio de Chebyshev de primera especie y se designa con $T_n(x)$.

DEFINICIÓN. *El polinomio de Chebyshev* $T_n(x)$ *se define para todo* x *real por medio de la ecuación*

$$T_n(x) = \sum_{k=0}^{[n/2]} \binom{n}{2k} x^{n-2k}(x^2 - 1)^k.$$

De la igualdad (15.28) obtenemos el teorema siguiente.

TEOREMA 15.7. *Si* $-1 \leq x \leq 1$ *entonces tenemos*

$$T_n(x) = \cos(n \operatorname{arc} \cos x).$$

Demostración. Si $\theta = \operatorname{arc} \cos x$ entonces $x = \cos \theta$ y $T_n(x) = \cos n\theta$.

Los polinomios de Chebyshev pueden calcularse rápidamente tomando la parte real de $(x + iy)^n$ poniendo $y^2 = 1 - x^2$, o utilizando la siguiente fórmula de recurrencia.

TEOREMA 15.8. *Los polinomios de Chebyshev satisfacen la siguiente fórmula de recurrencia*

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \qquad para\ n \geq 1,$$

siendo $T_0(x) = 1$ *y* $T_1(x) = x$.

Demostración. Supongamos primero que $-1 \leq x \leq 1$ y pongamos $x = \cos \theta$ en la identidad trigonométrica

$$\cos(n + 1)\theta + \cos(n - 1)\theta = 2 \cos \theta \cos n\theta.$$

Esto demuestra que $T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x)$ para x en el intervalo $-1 \leq$ $\leq x \leq 1$. Pero ya que ambos miembros son polinomios, esta relación debe ser válida para todo x.

Los cinco polinomios siguientes son

$$T_2(x) = 2x^2 - 1, \qquad T_3(x) = 4x^3 - 3x, \qquad T_4(x) = 8x^4 - 8x^2 + 1,$$
$$T_5(x) = 16x^5 - 20x^3 + 5x, \qquad T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1.$$

La fórmula de recurrencia demuestra que todos los coeficientes de $T_n(x)$ son enteros; además, el coeficiente de x^n es 2^{n-1}.

El teorema que sigue demuestra que $T_n(x)$ tiene exactamente n ceros de primer orden situados todos en el intervalo $[-1, 1]$.

TEOREMA 15.9. *Si $n \geq 1$ el polinomio $T_n(x)$ tiene ceros en los n puntos*

$$x_k = \cos \frac{(2k+1)\pi}{2n}, \qquad k = 0, 1, 2, \ldots, n-1.$$

Luego $T_n(x)$ admite la factorización

$$T_n(x) = 2^{n-1}(x - x_0)(x - x_1) \cdots (x - x_{n-1}) = 2^{n-1} \prod_{k=0}^{n-1} \left(x - \cos \frac{(2k+1)\pi}{2n} \right).$$

Demostración. Usamos la fórmula $T_n(x) = \cos n\theta$. Puesto que $\cos n\theta = 0$ sólo si $n\theta$ es un múltiplo impar de $\pi/2$, tenemos $T_n(x) = 0$ para x en $[-1, 1]$ únicamente si $n \text{ arc cos } x = (2k+1)\pi/2$ para un cierto entero k. Por tanto, los ceros de T_n en el intervalo $[-1, 1]$ se encuentran entre los números

$$x_k = \cos \frac{2k+1}{n} \frac{\pi}{2}, \qquad k = 0, \pm 1, \pm 2, \ldots. \tag{15.29}$$

Los valores $k = 0, 1, 2, \ldots, n-1$ dan n ceros distintos $x_0, x_1, \ldots, x_{n-1}$, situados todos en el intervalo abierto $(-1, 1)$. Puesto que un polinomio de grado n no puede tener más de n ceros, ésos deben ser *todos* los ceros de T_n. Los restantes x_k de (15.29) son repeticiones de esos n.

TEOREMA 15.10. *En el intervalo $[-1, 1]$ los valores extremos de $T_n(x)$ son $+1$ y -1, alcanzados alternativamente en los $n + 1$ puntos*

$$t_k = \cos \frac{k\pi}{n}, \qquad \textit{para } k = 0, 1, 2, \ldots, n. \tag{15.30}$$

Demostración. Según el teorema de Rolle, el máximo y el mínimo relativos de T_n deben presentarse entre dos ceros consecutivos; existen $n - 1$ puntos

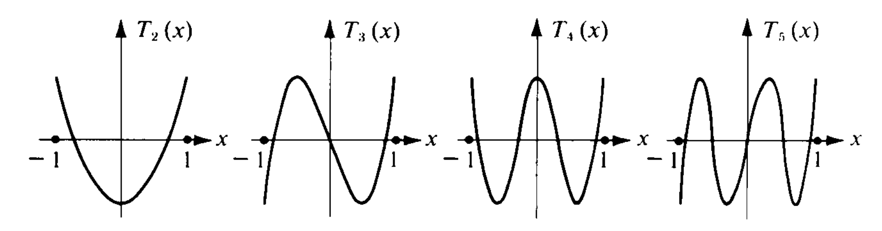


FIGURA 15.2 *Gráficas de polinomios de Chebyshev en el intervalo* $[-1, 1]$.

de tal naturaleza en el intervalo abierto $(-1, 1)$. En la fórmula del coseno para T_n vemos que los valores extremos, ± 1, son alcanzados en los $n-1$ puntos interiores $\cos(k\pi/n)$, $k = 1, 2, \ldots, n-1$, y también en los extremos $x = 1$ y $x = -1$. Por consiguiente, en el intervalo cerrado $[-1, 1]$ los valores extremos $+1$ y -1 son alcanzados *alternativamente* en los $n+1$ puntos $t_0, t_1, \ldots, t_n$ dados por $t_k = \cos(k\pi/n)$ para $k = 0, 1, 2, \ldots, n$.

La figura 15.2 representa las gráficas de $T_2, \ldots, T_5$ en el intervalo $[-1, 1]$.

15.16 Propiedad de mínimo de los polinomios de Chebyshev

Volvamos ahora al problema de encontrar un polinomio de grado prefijado para el que la norma del máximo sea lo más pequeña posible. El problema se resuelve con el siguiente teorema.

TEOREMA 15.11. *Sea $p_n(x) = x^n + \ldots$ un polinomio cualquiera de grado $n \geq 1$ con primer coeficiente igual a 1, y pongamos*

$$\|p_n\| = \max_{-1 \leq x \leq 1} |p_n(x)|.$$

Tenemos entonces la desigualdad

$$\|p_n\| \geq \|\tilde{T}_n\|, \tag{15.31}$$

en donde $\tilde{T}_n(x) = T_n(x)/2^{n-1}$. Además, la igualdad en (15.31) *es válida si $p_n = \tilde{T}_n$.*

Demostración. En el intervalo $[-1, 1]$ el polinomio $\tilde{T}_n$ toma sus valores extremos, $1/2^{n-1}$ y $-1/2^{n-1}$, alternativamente en los $n+1$ puntos distintos t_k de la ecuación (15.30). Por consiguiente, $\|\tilde{T}_n\| = 1/2^{n-1}$.

Demostremos a continuación que la desigualdad

$$\|p_n\| < \frac{1}{2^{n-1}} \tag{15.32}$$

nos lleva a una contradicción. Supongamos que p_n satisfaga (15.32) y consideremos la diferencia

$$r(x) = \tilde{T}_n(x) - p_n(x).$$

En los puntos t_k dados por (15.30) tenemos

$$r(t_k) = \frac{(-1)^k}{2^{n-1}} - p_n(t_k) = (-1)^k\left[\frac{1}{2^{n-1}} - (-1)^k p_n(t_k)\right].$$

Debido a (15.32) el factor entre corchetes es positivo. Por consiguiente, $r(t_k)$ tiene signos alternados en los $n + 1$ puntos $t_0, t_1, \ldots, t_n$. Puesto que r es continua debe anularse por lo menos una vez entre dos cambios de signo consecutivos. Por consiguiente, r tiene por lo menos n ceros distintos. Pero como r es un polinomio de grado $\leq n - 1$, esto significa que r es idénticamente nula. Por lo tanto, $P_n = \tilde{T}_n$, así que $\|P_n\| = \|\tilde{T}_n\| = 1/2^{n-1}$, en contradicción con (15.32). Esto demuestra que debe ser $\|p_n\| \geq 1/2^{n-1} = \|\tilde{T}_n\|$.

Aunque el teorema 15.11 hace referencia al intervalo $[-1, 1]$ y a un polinomio con primer coeficiente 1, puede también utilizarse para deducir un resultado análogo para un intervalo cualquiera $[a, b]$ y un polinomio cualquiera.

TEOREMA 15.12. *Sea* $q_n(x) = c_n x^n + \ldots$ *un polinomio de grado* $n \geq 1$, *y pongamos*

$$\|q_n\| = \max_{a \leq x \leq b} |q_n(x)|.$$

Tenemos entonces la desigualdad

$$\|q_n\| \geq |c_n| \frac{(b - a)^n}{2^{2n-1}}. \tag{15.33}$$

Además, la igualdad en (15.33) *es válida si*

$$q_n(x) = c_n \frac{(b - a)^n}{2^{2n-1}} T_n\left(\frac{2x - a - b}{b - a}\right).$$

Demostración. Consideremos la transformación

$$t = \frac{2x - a - b}{b - a}.$$

Ésta aplica el intervalo $a \leq x \leq b$ en forma uno a uno sobre el intervalo $-1 \leq t \leq 1$. Puesto que

$$x = \frac{b - a}{2} t + \frac{b + a}{2}$$

tenemos

$$x^n = \left(\frac{b - a}{2}\right)^n t^n + \text{términos de grado inferior}$$

luego

$$q_n(x) = c_n \left(\frac{b - a}{2}\right)^n p_n(t),$$

en donde $p_n(t)$ es un polinomio en t de grado n con primer coeficiente 1. Aplicando el teorema 15.11 a p_n obtenemos el teorema 15.12.

15.17 Aplicación a la fórmula del error en la interpolación

Volvamos ahora a la fórmula (15.27) del error en la interpolación por polinomios. Si elegimos los puntos de interpolación $x_0, x_1, \ldots, x_n$ de modo que coincidan con los $n + 1$ ceros del polinomio de Chebyshev T_{n+1} podemos escribir (15.27) en la forma

$$f(x) - P(x) = \frac{T_{n+1}(x)}{2^n(n+1)!} f^{(n+1)}(c).$$

Los puntos $x_0, x_1, \ldots, x_n$ están todos en el intervalo abierto $(-1, 1)$ y vienen dados por

$$x_k = \cos\left(\frac{2k+1}{n+1}\frac{\pi}{2}\right) \quad \text{para} \quad k = 0, 1, 2, \ldots, n.$$

Si x está en el intervalo $[-1, 1]$ tenemos $|T_{n+1}(x)| \leq 1$ y el error viene estimado por la desigualdad

$$|f(x) - P(x)| \leq \frac{1}{2^n(n+1)!} |f^{(n+1)}(c)|.$$

Si la interpolación se efectúa en el intervalo $[a, b]$ con los puntos

$$y_k = \frac{b-a}{2} x_k + \frac{b+a}{2}$$

como puntos de interpolación, el producto

$$A(x) = (x - y_0)(x - y_1) \cdots (x - y_n)$$

satisface la desigualdad $|A(x)| \leq (b-a)^{n+1}/2^{2n+1}$ para todo x en $[a, b]$. La estimación correspondiente para $f(x) - P(x)$ es

$$|f(x) - P(x)| \leq \frac{(b-a)^{n+1}}{2^{2n+1}(n+1)!} |f^{(n+1)}(c)|.$$

15.18 Ejercicios

En este conjunto de ejercicios T_n representa el polinomio de Chebyshev de grado n.

1. Demostrar que $T_n(-x) = (-1)^n T_n(x)$. Esto demuestra que T_n es función par si n es par y función impar si n es impar.

2. a) Demostrar que en el intervalo abierto $(-1, 1)$ la derivada T'_n viene dada por la fórmula

$$T_n'(x) = \frac{n \operatorname{sen} n\theta}{\operatorname{sen} \theta}, \qquad \text{donde} \quad \theta = \operatorname{arc} \cos x.$$

b) Calcular $T_n'(1)$ y $T_n'(-1)$.

3. Demostrar que $\displaystyle\int_0^x T_n(u)\, du = \frac{1}{2}\left\{\frac{T_{n+1}(x) - T_{n+1}(0)}{n+1} - \frac{T_{n-1}(x) - T_{n-1}(0)}{n-1}\right\}$ si $n \geq 2$.

4. a) Demostrar que $2T_m(x)T_n(x) = T_{m+n}(x) + T_{m-n}(x)$.
b) Demostrar que $T_{mn}(x) = T_m[T_n(x)] = T_n[T_m(x)]$.
5. Si $x = \cos \theta$, demostrar que $\operatorname{sen} \theta \operatorname{sen} n\theta$ es un polinomio en x, y determinar su grado.
6. El polinomio de Chebyshev T_n satisface la ecuación diferencial

$$(1 - x^2)y'' - xy' + n^2y = 0$$

en todo el eje real. Demostrarlo con cada uno de los métodos siguientes:
a) Derivando la relación $T_n'(x) \operatorname{sen} \theta = n \operatorname{sen} n\theta$ obtenida en el ejercicio 2 a).
b) Introducir el cambio de variable $x = \cos \theta$ en la ecuación diferencial

$$\frac{d^2 (\cos n\theta)}{d\theta^2} = -n^2 \cos n\theta.$$

7. Determinar, en función de polinomios de Chebyshev, un polinomio $Q(x)$ de grado $\leq n$ de aproximación óptima para x^{n+1} en el intervalo $[-1, 1]$ relativo a la norma del máximo.
8. Hallar un polinomio de grado ≤ 4 de aproximación óptima para la función $f(x) = x^5$ en el intervalo $[0, 1]$, relativo a la norma del máximo.
9. Un polinomio P se llama *primario* o *monario* si el coeficiente del término de mayor grado es 1. Para un intervalo dado $[a, b]$ designemos con $\|P\|$ el máximo de $|P|$ en $[a, b]$. Demostrar cada una de las siguientes afirmaciones:
a) Si $b - a < 4$, para todo $\epsilon > 0$ existe un polinomio monario P tal que $\|P\| < \epsilon$.
b) Si para todo $\epsilon > 0$ existe un polinomio monario P tal que $\|P\| < \epsilon$, entonces $b - a < 4$. Dicho de otro modo, los polinomios monarios con norma arbitrariamente pequeña existen si, y sólo si, el intervalo $[a, b]$ tiene longitud menor que 4.
10. Los polinomios de Chebyshev satisfacen las siguientes relaciones de ortogonalidad:

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1 - x^2}}\, dx = \begin{cases} 0 & \text{si} \quad n \neq m, \\ \pi & \text{si} \quad n = m = 0, \\ \dfrac{\pi}{2} & \text{si} \quad n = m > 0. \end{cases}$$

Demostrarlas por cada uno de los siguientes métodos:

a) A partir de la ecuación diferencial del ejercicio 6 deducir que

$$T_m(x)\frac{d}{dx}\left(\sqrt{1-x^2}\,T_n'(x)\right)+n^2\frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}=0.$$

Escribir la fórmula correspondiente intercambiando n y m, restar las dos ecuaciones, e integrar entre -1 y 1.

b) Utilizar las relaciones de ortogonalidad

$$\int_0^\pi \cos m\theta \cos n\theta\, d\theta = \begin{cases} 0 & \text{si} \quad n \neq m, \quad n > 0, \quad m > 0, \\ \pi & \text{si} \quad n = m = 0, \\ \dfrac{\pi}{2} & \text{si} \quad n = m > 0, \end{cases}$$

e introducir el cambio de variable $x = \cos\theta$.

11. Demostrar que para $-1 < x < 1$ tenemos

$$\frac{T_n(x)}{\sqrt{1-x^2}} = (-1)^n \frac{2^n n!}{(2n)!}\frac{d^n}{dx^n}(1-x^2)^{n-1/2}.$$

12. Sean $y_1, y_2, \ldots, y_n$ n números reales, y

$$x_k = \cos\frac{(2k-1)\pi}{2n} \qquad \text{para} \quad k = 1, 2, \ldots, n.$$

Sea P el polinomio de grado $\leq n-1$ que toma el valor y_k en x_k para $1 \leq k \leq n$. Si x no es uno de esos x_k demostrar que

$$P(x) = \frac{1}{n}\sum_{k=1}^{n}(-1)^{k-1}y_k\sqrt{1-x_k^2}\,\frac{T_n(x)}{x-x_k}.$$

13. Sea P un polinomio de grado $\leq n-1$ tal que

$$\sqrt{1-x^2}\,|P(x)| \leq 1$$

para $-1 \leq x \leq 1$. Demostrar que $\|P\| \leq n$, siendo $\|P\|$ el máximo de $|P|$ en el intervalo $[-1, 1]$. [*Indicación:* Utilizar el ejercicio 12. Considerar tres casos: $x_1 \leq x \leq 1$; $-1 \leq x \leq x_n$; $x_n \leq x \leq x_1$; en los dos primeros hacer uso del ejercicio 15 a) de la sección 15.9. En el tercer caso observar que $\sqrt{1-x^2} \geq \operatorname{sen}(\pi/2n) > 1/n$.]

En los ejercicios del 14 al 18, $U_n(x) = T_{n+1}'(x)/(n+1)$ para $n = 0, 1, 2, \ldots$.

14. a) Demostrar que $U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x)$ para $n \geq 2$.

b) Determinar la forma explícita de los polinomios $U_0, U_1, \ldots, U_5$.
c) Demostrar que $|U_n(x)| \leq n + 1$ si $-1 \leq x \leq 1$.

15. Demostrar que U_n satisface la ecuación diferencial

$$(1 - x^2)y'' - 3xy' + n(n + 2)y = 0.$$

16. Deducir las relaciones de ortogonalidad

$$\int_{-1}^{1} \sqrt{1 - x^2}\, U_m(x) U_n(x)\, dx = \begin{cases} 0 & \text{si} \quad m \neq n, \\ \dfrac{\pi}{2} & \text{si} \quad m = n. \end{cases}$$

17. Probar que

$$\sqrt{1 - x^2}\, U_n(x) = (-1)^n \frac{2^n (n + 1)!}{(2n + 1)!} \frac{d^n}{dx^n} (1 - x^2)^{n+1/2}$$

18. Sean $y_1, y_2, \ldots, y_n$ n números reales y

$$x_k = \cos \frac{k\pi}{n + 1} \qquad \text{para} \quad k = 1, 2, \ldots, n.$$

Sea P el polinomio de grado $\leq n - 1$ que toma el valor y_k en x_k para $1 \leq k \leq n$. Si x no es uno de los x_k demostrar que

$$P(x) = \frac{1}{n + 1} \sum_{k=1}^{n} (-1)^{k-1} (1 - x_k^2) y_k \frac{U_n(x)}{x - x_k}.$$

15.19 Integración aproximada. Regla de los trapecios

Muchos problemas tanto de Matemática pura como aplicada nos conducen a nuevas funciones cuyas propiedades no han sido estudiadas o cuyos valores no han sido tabulados. Con frecuencia es necesario obtener información cuantitativa sobre tales funciones, bien en forma gráfica, bien en forma numérica. Muchas de esas funciones se presentan como integrales del tipo

$$F(x) = \int_a^x f(t)\, dt,$$

en las que el integrando f viene dado por una fórmula analítica explícita o es conocida en parte por datos tabulados. El resto del presente capítulo se dedica a algunos de los más elementales métodos para encontrar aproximaciones numéricas de tales integrales. La idea básica es muy sencilla. Aproximamos el integrando f

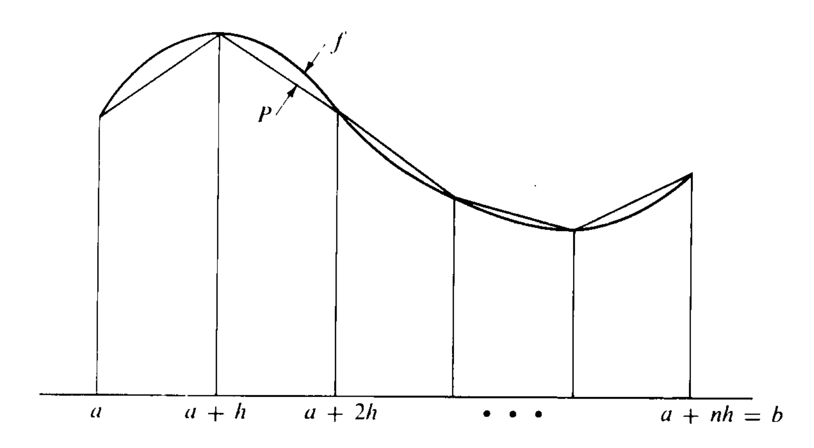


FIGURA 15.3 *Regla de los trapecios obtenida mediante interpolación lineal*

mediante otra función P cuya integral es de fácil cálculo, y tomamos entonces la integral de P como aproximación de la integral de f.

Si f es no negativa la integral $\int_a^b f(x)dx$ representa el área del conjunto de ordenadas de f sobre $[a, b]$. Esta interpretación geométrica de la integral sugiere inmediatamente ciertos procedimientos de integración aproximada. La figura 15.3 muestra un ejemplo de una función f con valores conocidos en $n + 1$ puntos igualmente separados $a, a + h, a + 2h, \ldots, a + nh = b$, donde $h = (b - a)/n$. Sea $x_k = a + kh$. Para cada $k = 0, 1, 2, \ldots, n - 1$ la gráfica de f correspondiente al intervalo $[x_k, x_{k+1}]$ ha sido aproximada mediante una función lineal que coincida con f en los extremos x_k y x_{k+1}. Con P designemos la correspondiente función poligonal de interpolación definida sobre todo el intervalo $[a, b]$. Tenemos entonces

$$(15.34) \qquad P(x) = \frac{x_{k+1} - x}{h} f(x_k) + \frac{x - x_k}{h} f(x_{k+1}) \qquad \text{si } \; x_k \leq x \leq x_{k+1}.$$

Integrando en el intervalo $[x_k, \; x_{k+1}]$ encontramos que

$$\int_{x_k}^{x_{k+1}} P(x)\, dx = h \frac{f(x_k) + f(x_{k+1})}{2}.$$

Cuando f es positiva ésta es el área de un trapecio determinado por la gráfica de P sobre $[x_k, x_{k+1}]$. La fórmula es válida, naturalmente, incluso si f no es siem-

pre positiva. Sumando las integrales correspondientes a todos los subintervalos $[x_k, x_{k+1}]$ obtenemos

$$\int_a^b P(x)\,dx = \frac{h}{2}\sum_{k=0}^{n-1}[f(x_k) + f(x_{k+1})] \tag{15.35}$$

$$= \frac{h}{2}\left(f(a) + 2\sum_{k=1}^{n-1} f(a + kh) + f(b)\right).$$

Para utilizar esta suma como una aproximación de la integral $\int_a^b f(x)dx$ necesitamos una estimación del error $\int_a^b f(x)dx - \int_a^b P(x)dx$. Si f posee derivada segunda continua en $[a, b]$ este error viene dado por el siguiente teorema.

TEOREMA 15.13. REGLA DE LOS TRAPECIOS. *Si suponemos que f tiene derivada segunda f'' continua en $[a, b]$ y si n es un número natural, pongamos $h=(b-a)/n$, entonces tenemos*

$$\int_a^b f(x)\,dx = \frac{b-a}{2n}\left(f(a) + 2\sum_{k=1}^{n-1} f(a + kh) + f(b)\right) - \frac{(b-a)^3}{12n^2}f''(c) \tag{15.36}$$

para un cierto c en $[a, b]$.

Observación. La fórmula (15.36) es la *regla de los trapecios.* El término $-f''(c)(b-a)^3/12n^2$ representa el error al aproximar $\int_a^b f(x)\,dx$ por la $\int_a^b P(x)\,dx$. En cuanto se conoce el valor máximo de f'' en $[a, b]$ podemos aproximar la integral de f con el grado de precisión que se desee sin más que tomar n suficientemente grande. Observemos que no se requiere el conocimiento de la función P para usar esa fórmula. Tan sólo necesitamos conocer los valores de f en los puntos $a, a+h, \ldots, a+nh$, y tener una estimación de $|f''(c)|$.

Demostración. Sea P la función de interpolación dada por (15.34), donde $x_k = a + kh$. En cada subintervalo $[x_k, x_{k+1}]$ aplicamos la estima del error para la interpolación lineal dada por el teorema 15.3 y encontramos

$$f(x) - P(x) = (x - x_k)(x - x_{k+1})\frac{f''(c_k)}{2!} \tag{15.37}$$

para un cierto c_k en (x_k, x_{k+1}). Sean M_2 y m_2 el máximo y el mínimo, respectivamente, de f'' en $[a, b]$, y sea

$$B(x) = (x - x_k)(x_{k+1} - x)/2.$$

Entonces $B(x) \geq 0$ en el intervalo $[x_k, x_{k+1}]$, y de (15.37) obtenemos las desigualdades

$$m_2B(x) \leq P(x) - f(x) \leq M_2B(x)$$

en este intervalo. Integrando, tenemos

$$(15.38) \qquad m_2 \int_{x_k}^{x_k+h} B(x)\,dx \le \int_{x_k}^{x_k+h} [P(x) - f(x)]\,dx \le M_2 \int_{x_k}^{x_k+h} B(x)\,dx\,.$$

La integral de B viene dada por

$$\int_{x_k}^{x_k+h} B(x)\,dx = \frac{1}{2}\int_{x_k}^{x_k+h} (x - x_k)(x_{k+1} - x)\,dx = \frac{1}{2}\int_0^h t(h-t)\,dt = \frac{h^3}{12}\,.$$

Por consiguiente, las desigualdades (15.38) nos dan

$$m_2 \le \frac{12}{h^3}\int_{x_k}^{x_k+h} [P(x) - f(x)]\,dx \le M_2\,.$$

Sumando esas desigualdades para $k = 0, 1, 2, \ldots, n-1$ y dividiendo por n, obtenemos

$$m_2 \le \frac{12}{nh^3}\int_a^b [P(x) - f(x)]\,dx \le M_2\,.$$

Puesto que la función f'' es continua en $[a, b]$, alcanza cualquier valor comprendido entre su mínimo m_2 y su máximo M_2 en algún punto de $[a, b]$. En particular, tenemos

$$f''(c) = \frac{12}{nh^3}\int_a^b [P(x) - f(x)]\,dx$$

para algún c en $[a, b]$. Dicho de otro modo,

$$\int_a^b f(x)\,dx = \int_a^b P(x)\,dx - \frac{nh^3}{12} f''(c).$$

Con (15.35) y la relación $h = (b - a)/n$ obtenemos (15.36).

Para deducir la regla de los trapecios utilizamos un polinomio lineal para interpolar entre cada par de valores consecutivos de f. Se pueden obtener fórmulas más precisas interpolando con polinomios de grado superior. En la próxima sección consideramos un caso particular importante notable por su simplicidad y precisión.

15.20 Regla de Simpson

La curva de trazo continuo de la figura 15.4 es la gráfica de una función f correspondiente al intervalo $[a, b]$. El punto medio del intervalo, $(a + b)/2$, se

representa con m. La curva de puntos es la gráfica de un polinomio cuadrático P que coincide con f en los tres puntos a, m y b. Si utilizamos la integral $\int_a^b P(x)dx$ como una aproximación de $\int_a^b f(x)dx$ vamos a parar a una fórmula de integración aproximada conocida con el nombre de *regla de Simpson.*

En lugar de determinar P explícitamente, introducimos una transformación lineal que lleva el intervalo $[a, b]$ sobre el intervalo $[0, 2]$. Si escribimos

$$t = \frac{x - a}{m - a}, \qquad \text{o} \qquad x = a + (m - a)t,$$

vemos que t toma los valores 0, 1, 2 cuando x toma los a, m, b. Sea ahora

$$\varphi(t) = P[a + (m - a)t].$$

Entonces $\varphi(t)$ es un polinomio cuadrático en t que toma los valores $P(a)$, $P(m)$, $P(b)$ en los puntos $t = 0, 1, 2$, respectivamente. También, tenemos

$$\int_0^2 \varphi(t)\, dt = \int_0^2 P[a + (m - a)t]\, dt = \frac{1}{m - a} \int_a^b P(x)\, dx\,;$$

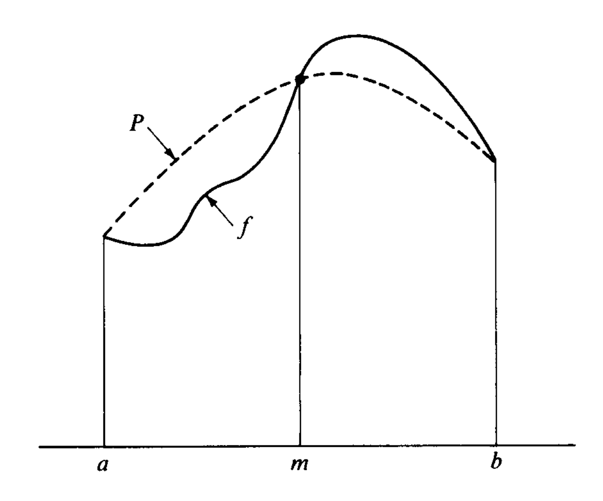


FIGURA 15.4 *Interpolación mediante un polinomio cuadrático P.*

luego

$$\text{(15.39)} \qquad \int_a^b P(x)\, dx = (m - a) \int_0^2 \varphi(t)\, dt = \frac{b - a}{2} \int_0^2 \varphi(t)\, dt\,.$$

Utilicemos ahora la fórmula de interpolación de Newton para construir φ. Tenemos

$$\varphi(t) = \varphi(0) + t\,\Delta\varphi(0) + t(t-1)\frac{\Delta^2\varphi(0)}{2!},$$

donde $\Delta\varphi(t) = \varphi(t+1) - \varphi(t)$. Integrando entre 0 y 2 obtenemos

$$\int_0^2 \varphi(t)\,dt = 2\varphi(0) + 2\,\Delta\varphi(0) + \tfrac{1}{3}\Delta^2\varphi(0).$$

Puesto que $\Delta\varphi(0) = \varphi(1) - \varphi(0)$ y $\Delta^2\varphi(0) = \varphi(2) - 2\varphi(1) + \varphi(0)$, la integral es igual a

$$\int_0^2 \varphi(t)\,dt = \tfrac{1}{3}[\varphi(0) + 4\varphi(1) + \varphi(2)] = \tfrac{1}{3}[P(a) + 4P(m) + P(b)].$$

Teniendo en cuenta (15.39) y que P coincide con f en a, m, b, obtenemos

(15.40) $$\int_a^b P(x)\,dx = \frac{b-a}{6}[f(a) + 4f(m) + f(b)].$$

Por consiguiente, podemos escribir

$$\int_a^b f(x)\,dx = \frac{b-a}{6}[f(a) + 4f(m) + f(b)] + R,$$

siendo $R = \int_a^b f(x)\,dx - \int_a^b P(x)\,dx$.

Si f es polinomio cuadrático, P es idéntico a f y el error R es cero. También se observa que $R = 0$ cuando f es un polinomio *cúbico*. Para demostrarlo utilizamos la estimación de error para la interpolación de Lagrange dada por el teorema 15.3, y escribimos

(15.41) $$f(x) - P(x) = (x-a)(x-m)(x-b)\frac{f'''(c)}{3!},$$

donde $c \in (a, b)$. Cuando f es un polinomio cúbico la derivada tercera es constante, $f'''(x) = C$, y la fórmula anterior se transforma en

$$f(x) - P(x) = \frac{C}{6}(x-a)(x-m)(x-b) = \frac{C}{6}(t+h)t(t-h),$$

donde $t = x - m$ y $h = (b-a)/2$. Por tanto

$$R = \int_a^b [f(x) - P(x)]\,dx = \frac{C}{6}\int_{-h}^{h}(t^3 - h^2 t)\,dt = 0,$$

ya que la última integral es una función impar. Esta propiedad queda interpretada en la figura 15.5. La curva punteada es la gráfica de un polinomio cúbico f que coincide en P en a, m, b. En este caso $R = \int_a^b [f(x) - P(x)]dx = A_1 - A_2$, donde A_1 y A_2 son las áreas de las dos regiones sombreadas. Como que $R = 0$ las dos regiones tienen áreas iguales.

Acabamos de ver que la ecuación (15.40) es válida si P es un polinomio de grado ≤ 3 que coincide con f en a, m y b. Eligiendo cuidadosamente este polinomio podemos mejorar considerablemente la estimación del error (15.41). Ya hemos impuesto tres condiciones a P, a saber, $P(a) = f(a)$, $P(m) = f(m)$, $P(b) = f(b)$. Imponemos ahora una cuarta condición, $P'(m) = f'(m)$. Esto hace que P y f tengan la misma pendiente en $(m, f(m))$, y podemos esperar que esto mejorará la aproximación de f por P en el intervalo $[a, b]$.

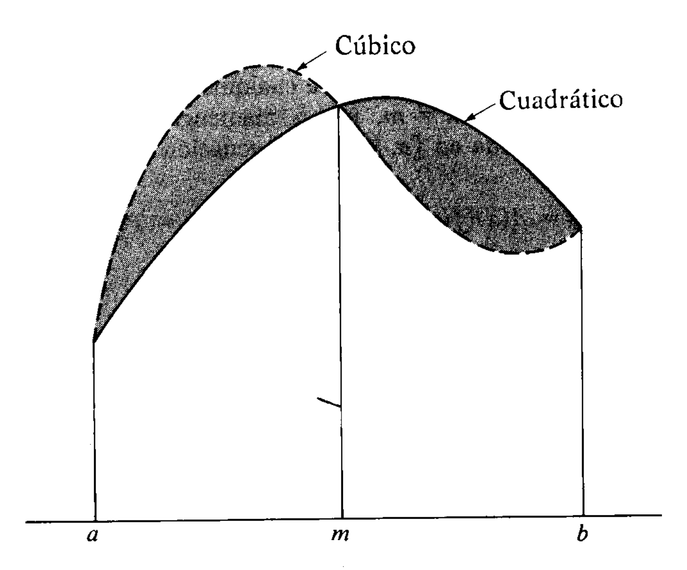


FIGURA 15.5 *Las dos regiones sombreadas tienen la misma área para cualquier polinomio cúbico de interpolación.*

Para demostrar que un tal P puede elegirse siempre, sea Q el polinomio cuadrático que coincide con f en a, m, b, y sea

$$P(x) = Q(x) + A(x - a)(x - m)(x - b),$$

donde A es una constante que hay que determinar. Cualquiera que sea A, este polinomio cúbico P coincide con Q y por tanto con f en a, m, b. Elijamos ahora A

de modo que $P'(m) = f'(m)$. Derivando la fórmula anterior y poniendo $x = m$ obtenemos

$$P'(m) = Q'(m) + A(m - a)(m - b).$$

Por consiguiente, si tomamos $A = [f'(m) - Q'(m)]/[(m - a)(m - b)]$ se satisface también la condición $P'(m) = f'(m)$.

A continuación demostramos que para este P elegido tenemos

$$f(x) - P(x) = (x - a)(x - m)^2(x - b)\frac{f^{(4)}(z)}{4!} \tag{15.42}$$

para algún z en (a, b), con tal de que exista la cuarta derivada $f^{(4)}$ en $[a, b]$. Para demostrar (15.42) razonamos como en la demostración del teorema 15.3. Observamos primero que (15.42) se satisface para cualquier z si $x = a$, m o b. Por consiguiente, supuesto que $x \neq a$, $x \neq m$, $x \neq b$, mantengamos x fijo, e introducimos una nueva función F definida en $[a, b]$ por la ecuación

$$F(t) = A(x)[f(t) - P(t)] - A(t)[f(x) - P(x)],$$

donde

$$A(t) = (t - a)(t - m)^2(t - b).$$

Observemos que $F(t) = 0$ para $t = a, m, b$, y x. Según el teorema de Rolle, $F'(t)$ se anula en cada uno de los tres intervalos abiertos determinados por esos cuatro puntos. Además, $F'(m) = 0$ debido a que $A'(m) = 0$ y $f'(m) = P'(m)$. Por consiguiente $F'(t) = 0$ por lo menos en cuatro puntos distintos de (a, b). Por el teorema de Rolle $F''(t) = 0$ por lo menos en tres puntos, $F'''(t) = 0$ por lo menos en dos puntos, y $F^{(4)}(t) = 0$ por lo menos en un punto, sea éste $t = z$. De la definición de F resulta

$$F^{(4)}(t) = A(x)[f^{(4)}(t) - P^{(4)}(t)] - A^{(4)}(t)[f(x) - P(x)]$$

$$= A(x)f^{(4)}(t) - 4!\,[f(x) - P(x)].$$

Cuando sustituimos $t = z$ en esta ecuación obtenemos (15.42).

Es cosa sencilla ahora demostrar la regla de Simpson en la forma siguiente.

TEOREMA 15.14. REGLA DE SIMPSON. *Si f tiene derivada cuarta continua en $[a, b]$ y si $m = (a + b)/2$, entonces*

$$\int_a^b f(x)\,dx = \frac{b - a}{6}[f(a) + 4f(m) + f(b)] - \frac{(b - a)^5}{2880}f^{(4)}(c) \tag{15.43}$$

para un cierto c de $[a, b]$.

Demostración. Representemos con M_4 y m_4, respectivamente, los valores máximo y mínimo de $f^{(4)}$ en $[a, b]$, y sea $B(x) = -(x - a)(x - m)^2(x - b)/4!$. Ya que $B(x) \geq 0$ para cada x en $[a, b]$, la ecuación (15.42) nos lleva a las desigualdades

$$m_4 B(x) \leq P(x) - f(x) \leq M_4 B(x).$$

Integrando, encontramos

$$(15.44) \qquad m_4 \int_a^b B(x)\,dx \leq \int_a^b [P(x) - f(x)]\,dx \leq M_4 \int_a^b B(x)\,dx.$$

Para calcular la integral $\int_a^b B(x)dx$ tomamos $h = (b - a)/2$ y tenemos

$$\int_a^b B(x)\,dx = -\frac{1}{4!}\int_a^b (x - a)(x - m)^2(x - b)\,dx = -\frac{1}{4!}\int_{-h}^h (t + h)t^2(t - h)\,dt$$
$$= -\frac{2}{4!}\int_0^h t^2(t^2 - h^2)\,dt = \frac{1}{4!}\frac{4h^5}{15} = \frac{(b - a)^5}{2880}.$$

Por tanto, las desigualdades de (15.44) nos dan

$$m_4 \leq \frac{2880}{(b - a)^5}\int_a^b [P(x) - f(x)]\,dx \leq M_4.$$

Pero ya que $f^{(4)}$ es continua en $[a, b]$, alcanza un valor cualquiera comprendido entre su mínimo m_4 y su máximo M_4 en algún punto de $[a, b]$. Luego

$$f^{(4)}(c) = \frac{2880}{(b - a)^5}\int_a^b [P(x) - f(x)]\,dx$$

para algún c en $[a, b]$. Puesto que $\int_a^b P(x)dx = \frac{1}{6}(b - a)[f(a) + 4f(m) + f(b)]$, esta ecuación nos da (15.43).

La regla de Simpson es de especial interés porque su precisión es mayor de lo que podría esperarse a partir del conocimiento de la función f tan sólo en tres puntos. Si los valores de f se conocen en un número impar de puntos igualmente separados, $a, a + h, \ldots, a + 2nh$, ordinariamente es más sencillo aplicar la regla de Simpson sucesivamente a cada uno de los intervalos $[a, a + 2h]$, $[a + 2h, a + 4h]$, ..., que actuar con un polinomio de grado $\leq 2n$ sobre el intervalo completo $[a, a + 2nh]$. Aplicando la regla de Simpson de esta forma, obtenemos la siguiente extensión del teorema 15.14.

TEOREMA 15.15. REGLA DE SIMPSON GENERALIZADA. *Supongamos que f*

tenga derivada cuarta continua en $[a, b]$. *Sean* $h = (b-a)/(2n)$ *y* $f_k = f(a+kh)$ *para* $k = 1, 2, \ldots, 2n-1$. *Tenemos entonces*

$$\int_a^b f(x)\,dx = \frac{b-a}{6n}\left(f(a) + 4\sum_{k=1}^{n} f_{2k-1} + 2\sum_{k=1}^{n-1} f_{2k} + f(b)\right) - \frac{(b-a)^5}{2880n^4} f^{(4)}(\bar{c})$$

para un cierto $\bar{c}$ *en* $[a, b]$.

En el ejercicio 9 de la próxima sección se propone la demostración de este teorema.

15.21 Ejercicios

1. a) Aplicar la regla de los trapecios con $n = 10$ para calcular un valor aproximado de la integral

$$\log 2 = \int_1^2 \frac{dx}{x}.$$

Obtener cotas superior e inferior del error. [Véase ejercicio 10 b) para comparar la precisión con la que se obtiene con la regla de Simpson.]
b) ¿Cuál es el menor valor de n que puede garantizar seis cifras exactas en el cálculo de $\log 2$ por este método?

2. a) Demostrar que existe un número positivo c en el intervalo $[0, 1]$ tal que la fórmula

$$\int_{-1}^{1} f(x)\,dx = f(c) + f(-c)$$

es exacta para todos los polinomios de grado ≤ 3.
b) Generalizar el resultado de la parte a) para un intervalo cualquiera. Esto es, demostrar que existen dos constantes c_1 y c_2 en $[a, b]$ tal que la fórmula

$$\int_a^b f(x)\,dx = \frac{b-a}{2}\,[f(c_1) + f(c_2)]$$

es exacta para todos los polinomios de grado ≤ 3. Expresar c_1 y c_2 en función de a y b.

3. a) Demostrar que existe una constante positiva c tal que la fórmula

$$\int_{-1/2}^{1/2} f(x)\,dx = \tfrac{1}{3}[f(-c) + f(0) + f(c)]$$

es exacta para todos los polinomios de grado ≤ 3.
b) Generalizar el resultado de la parte a) a un intervalo cualquiera. Esto es, demostrar que existen dos constantes c_1 y c_2 en $[a, b]$ tales que la fórmula

$$\int_a^b f(x)\,dx = \frac{b-a}{3}\left[f(c_1) + f\left(\frac{a+b}{2}\right) + f(c_2)\right]$$

es exacta para todos los polinomios de grado ≤ 3. Expresar c_1 y c_2 en función de a y b.

4. Demostrar que existen dos constantes positivas a y b tales que la fórmula

$$\int_0^{\infty} e^{-x} f(x)\,dx = \tfrac{1}{4}[af(b) + bf(a)]$$

es exacta para todos los polinomios de grado ≤ 3.

5. Demostrar que existe una constante positiva c tal que la fórmula

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)\,dx = \frac{\sqrt{\pi}}{6}[f(-c) + 4f(0) + f(c)]$$

es exacta para todos los polinomios de grado ≤ 5.

6. Sea P_n el polinomio de interpolación de grado $\leq n$ que coincide con f en $n + 1$ puntos distintos $x_0, x_1, \ldots, x_n$.

a) Demostrar que existen las constantes $A_0(n), A_1(n), \ldots, A_n(n)$, que dependen tan sólo de los números $x_0, x_1, \ldots, x_n$, a, y b, y no de f, tales que

$$\int_a^b P_n(x)\,dx = \sum_{k=0}^{n} A_k(n) f(x_k).$$

Los números $A_k(n)$ se llaman *pesos*. (Algunas veces números de Christoffel.)

b) Para un conjunto dado de puntos de interpolación distintos y un intervalo dado $[a, b]$, sean $W_0(n), W_1(n), \ldots, W_n(n)$, $n + 1$ constantes tales que la fórmula

$$\int_a^b f(x)\,dx = \sum_{k=0}^{n} W_k(n) f(x_k)$$

es exacta para todos los polinomios de grado $\leq n$. Demostrar que

$$\sum_{k=0}^{n} x_k^r W_k(n) = \frac{b^{r+1} - a^{r+1}}{r + 1} \qquad \text{para} \quad r = 0, 1, \ldots, n.$$

Éste es un sistema de $n + 1$ ecuaciones lineales que puede usarse para determinar los pesos. Puede demostrarse que este sistema tiene siempre solución única. También puede demostrarse que para una elección adecuada de puntos de interpolación es posible hacer todos los pesos iguales. Cuando los pesos son todos iguales la fórmula de integración se llama fórmula de integración de Chebyshev. Los ejercicios 2 y 3 son ejemplos de esa fórmula. El ejercicio siguiente demuestra que para una elección adecuada de puntos de interpolación la fórmula de integración que resulta es exacta para todos los polinomios de grado $\leq 2n + 1$.

7. En este ejercicio se pueden usar las propiedades de los polinomios de Legendre establecidas en las secciones 6.19 y 6.20. Sean $x_0, x_1, \ldots, x_n$ los ceros del polinomio de Legendre $P_{n+1}(x)$. Esos ceros son distintos y están en el intervalo $[-1,1]$. Sea $f(x)$ un polinomio en x cualquiera de grado $\leq 2n + 1$. Dividamos $f(x)$ por $P_{n+1}(x)$ y escribamos

$$f(x) = P_{n+1}(x)Q(x) + R(x),$$

siendo los polinomios Q y R de grado $\leq n$.

a) Demostrar que el polinomio R coincide con f en los ceros de P_{n+1} y que

$$\int_{-1}^{1} f(x)\,dx = \int_{-1}^{1} R(x)\,dx.$$

b) Demostrar que existen $n + 1$ pesos $W_0(n), \ldots, W_n(n)$ (independientes de f) tales que

$$\int_{-1}^{1} f(x)\,dx = \sum_{k=0}^{n} W_k(n) f(x_k).$$

Esto da una fórmula de integración con $n + 1$ puntos de interpolación que es exacta para todos los polinomios de grado $\leq 2n + 1$.

c) Hacer $n = 2$ y demostrar que la fórmula de la parte b) se convierte en

$$\int_{-1}^{1} f(x)\,dx = \tfrac{5}{9} f(-\sqrt{\tfrac{3}{5}}) + \tfrac{8}{9} f(0) + \tfrac{5}{9} f(\sqrt{\tfrac{3}{5}}).$$

Esta fórmula es exacta para todos los polinomios de grado ≤ 5.

d) Introducir una transformación lineal conveniente y escribir nuevamente la fórmula de la parte c) para un intervalo cualquiera $[a, b]$.

8. En este ejercicio se expone un método debido a Peano para deducir la fórmula que da el error en la regla de Simpson.

a) Con la integración por partes reiterada deducir la relación

$$\int u(t)v'''(t)\,dt = u(t)v''(t) - u'(t)v'(t) + u''(t)v(t) - \int g(t)\,dt,$$

donde $g(t) = u'''(t)v(t)$.

b) Suponer que φ posee derivada cuarta continua en el intervalo $[-1, 1]$. Tomar

$$v(t) = t(1 - t)^2/6, \qquad u(t) = \varphi(t) + \varphi(-t),$$

y usando la parte a) demostrar que

$$\int_{-1}^{1} \varphi(t)\,dt = \tfrac{1}{3}[\varphi(-1) + 4\varphi(0) + \varphi(1)] - \int_{0}^{1} g(t)\,dt.$$

Demostrar luego que $\int_0^1 g(t)\,dt = \varphi^{(4)}(c)/90$ para un cierto c en $[-1, 1]$.

c) Con una transformación lineal conveniente deducir el teorema 15.14 a partir del resultado de la parte b).

9. a) Sean $a_1, a_2, \ldots, a_n$ n números no negativos cuya suma es 1. Supongamos que φ es continua en un intervalo $[a, b]$. Si $c_1, c_2, \ldots, c_n$ son n puntos cualesquiera en $[a, b]$ (no necesariamente distintos), demostrar que existe por lo menos un punto c en $[a, b]$ tal que

$$\sum_{k=1}^{n} a_k \varphi(c_k) = \varphi(c).$$

[*Indicación:* Designar con M y m el máximo y el mínimo respectivamente de φ en $[a, b]$ y emplear la desigualdad $m \leq \varphi(c_k) \leq M$.]

b) Usar la parte a) y el teorema 15.14 para deducir la regla de Simpson generalizada que se dio en el teorema 15.15.

10. Calcular $\log 2$ a partir de la fórmula $\log 2 = \int_1^2 x^{-1}\, dx$ utilizando la regla de Simpson generalizada con a) $n = 2$; b) $n = 5$. Dar en cada caso cotas superior e inferior del error.

11. a) Sea $\varphi(t)$ un polinomio en t de grado ≤ 3. Expresar $\varphi(t)$ mediante la fórmula de interpolación de Newton e integrar para deducir la fórmula

$$\int_0^3 \varphi(t)\, dt = \tfrac{3}{8}[\varphi(0) + 3\varphi(1) + 3\varphi(2) + \varphi(3)].$$

b) Sea P el polinomio de interpolación de grado ≤ 3 que coincide con f en los puntos $a, a + h, a + 2h, a + 3h$, siendo $h > 0$. Con la parte a) demostrar que

$$\int_a^{a+3h} P(x)\, dx = \frac{3h}{8}\,[f(a) + 3f(a + h) + 3f(a + 2h) + f(a + 3h)].$$

c) Suponer que f posee derivada continua en $[a, b]$, y sea $h = (b - a)/3$. Demostrar que

$$\int_a^b f(x)\, dx = \frac{b - a}{8}\,[f(a) + 3f(a + h) + 3f(a + 2h) + f(b)] - \frac{(b - a)^5}{6480} f^{(4)}(c)$$

para algún c de $[a,b]$. Esta fórmula de integración aproximada se llama *regla de Cotes.*

d) Con la regla de Cotes calcular $\log 2 = \int_1^2 x^{-1}\, dx$ y dar cotas superior e inferior del error.

12. a) Utilizar la ecuación vectorial $\boldsymbol{r}(t) = a \operatorname{sen} t\, \boldsymbol{i} + b \cos t\, \boldsymbol{j}$ siendo $0 < b < a$, y demostrar que la longitud L de una elipse viene dada por la integral

$$L = 4a \int_0^{\pi/2} \sqrt{1 - k^2 \operatorname{sen}^2 t}\, dt,$$

donde $k = \sqrt{a^2 - b^2}/a$.

b) Probar que la regla de Simpson da la fórmula

$$L = \frac{\pi}{3}\,[a + b + \sqrt{8(a^2 + b^2)}] - \frac{a\pi^5}{23040} f^{(4)}(c)$$

para un cierto c en $[0, \pi/2]$, siendo $f(t) = \sqrt{1 - k^2 \operatorname{sen}^2 t}$.

15.22 Fórmula de sumación de Euler

Sea n un entero positivo. Cuando la fórmula de los trapecios (teorema 15.13) se aplica al intervalo $[0, n]$ toma la forma

$$\int_0^n f(x)\, dx = \sum_{k=0}^{n-1} f(k) + \tfrac{1}{2}(f(n) - f(0)) - \frac{f''(c)n}{12}$$

para un cierto c en $[0, n]$. Si f es un polinomio cuadrático entonces f'' es constante y por tanto $f''(c) = f''(0)$. En este caso la fórmula puede volverse a escribir en la forma

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)\,dx + \frac{f(0) + f(n)}{2} + \frac{f''(0)n}{12}. \tag{15.45}$$

Esta fórmula es exacta cuando f es un polinomio cualquiera de grado ≤ 2.

Euler descubrió una notable generalización de esa fórmula que es válida para cualquier función con derivada primera continua. Puede utilizarse para aproximar integrales mediante sumas o, lo que es más frecuente, para calcular o aproximar sumas por medio de integrales. Por este motivo corrientemente se la llama fórmula de «sumación» y no de integración. Puede establecerse del siguiente modo.

TEOREMA 15.16. FÓRMULA DE SUMACIÓN DE EULER. *Si f tiene derivada continua en $[0, n]$, entonces*

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)\,dx + \frac{f(0) + f(n)}{2} + \int_0^n (x - [x] - \tfrac{1}{2})f'(x)\,dx, \tag{15.46}$$

en donde $[x]$ es el mayor entero $\leq x$.

Demostración. La integración por partes nos da

$$\int_0^n (x - \tfrac{1}{2})f'(x)\,dx = (n - \tfrac{1}{2})f(n) + \tfrac{1}{2}f(0) - \int_0^n f(x)\,dx. \tag{15.47}$$

Consideremos ahora la integral $\int_0^n [x]f'(x)dx$ y la escribimos como una suma de integrales en cada una de las cuales $[x]$ tiene un valor fijo. Así pues, tenemos

$$\begin{aligned}
\int_0^n [x]f'(x)\,dx &= \sum_{r=0}^{n-1} \int_r^{r+1} [x]f'(x)\,dx = \sum_{r=0}^{n-1} r \int_r^{r+1} f'(x)\,dx \\
&= \sum_{r=0}^{n-1} r(f(r+1) - f(r)) = \sum_{r=0}^{n-1} rf(r+1) - \sum_{r=0}^{n-1} rf(r) \\
&= -\sum_{r=0}^{n-1} f(r+1) + \sum_{r=0}^{n-1} (r+1)f(r+1) - \sum_{r=0}^{n-1} rf(r) \\
&= -\sum_{k=1}^{n} f(k) + nf(n) = -\sum_{k=0}^{n} f(k) + f(0) + nf(n).
\end{aligned}$$

Restando ésta de la ecuación (15.47) obtenemos

$$\int_0^n (x - [x] - \tfrac{1}{2})f'(x)\,dx = \sum_{k=0}^{n} f(k) - \frac{f(0) + f(n)}{2} - \int_0^n f(x)\,dx,$$

que es equivalente a (15.46).

La última integral del segundo miembro de (15.46) puede ponerse en la forma

$$\int_0^n (x - [x] - \tfrac{1}{2})f'(x)\,dx = \int_0^n \varphi_1(x)f'(x)\,dx,$$

en donde φ_1 es la función definida por

$$\varphi_1(x) = \begin{cases} x - [x] - \frac{1}{2} & \text{si} \quad x \text{ no es entero} \\ 0 & \text{si} \quad x \text{ es entero} \end{cases}$$

La gráfica de φ_1 es la de la figura 15.6 a). Observemos que $\varphi_1(x + 1) = \varphi_1(x)$, lo cual significa que φ_1 es periódica con período 1. Asimismo, si $0 < x < 1$ tenemos $\varphi_1(x) = x - \frac{1}{2}$, así que $\int_0^1 \varphi_1(t)dt = 0$.

La figura 15.6 b) es la gráfica de φ_2, integral indefinida de φ_1, dada por

$$\varphi_2(x) = \int_0^x \varphi_1(t)\,dt.$$

Es fácil comprobar que φ_2 también es periódica con período 1. Además, tenemos

$$\varphi_2(x) = \frac{x(x-1)}{2} \quad \text{si} \quad 0 \le x \le 1.$$

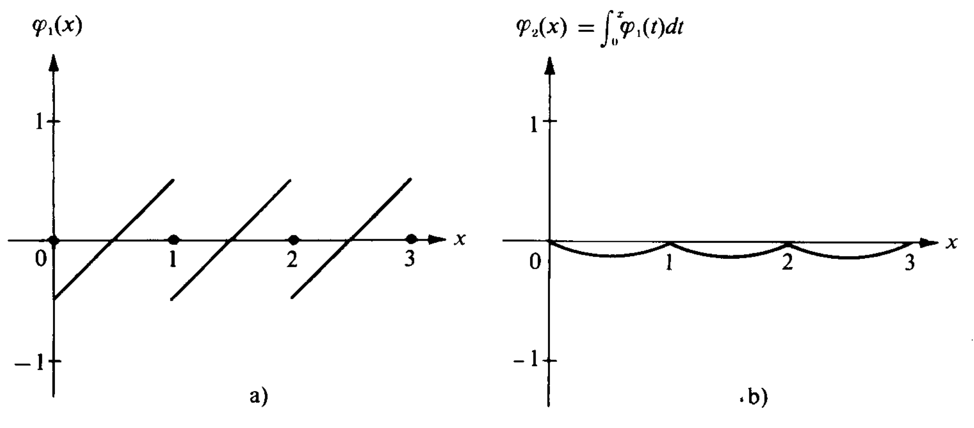


FIGURA 15.6 *Gráficas de las funciones periódicas* φ_1 *y* φ_2.

Esto prueba que $-\frac{1}{8} \leq \varphi_2(x) \leq 0$ para todo x. Las desigualdades $-\frac{1}{8} < \varphi_2(x) < 0$ son válidas salvo cuando x es un entero o la mitad de un entero.

El teorema siguiente da otra versión de la fórmula de sumación de Euler en función de φ_2.

TEOREMA 15.17. *Si f'' es continua en $[0, n]$ tenemos*

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)\,dx + \frac{f(0)+f(n)}{2} - \int_0^n \varphi_2(x) f''(x)\,dx. \tag{15.48}$$

Demostración. Puesto que $\varphi'_2(x) = \varphi_1(x)$ en los puntos de continuidad de φ_1, tenemos

$$\int_0^n \varphi_1(x) f'(x)\,dx = \int_0^n \varphi'_2(x) f'(x)\,dx.$$

La integración por partes nos da

$$\int_0^n \varphi'_2(x) f'(x)\,dx = \varphi_2(x) f'(x)\Big|_0^n - \int_0^n \varphi_2(x) f''(x)\,dx = -\int_0^n \varphi_2(x) f''(x)\,dx,$$

puesto que $\varphi_2(n) = \varphi_2(0) = 0$. Aplicando esta fórmula en (15.46) obtenemos (15.48).

Observación. Si bien los teoremas 15.16 y 15.17 se refieren al intervalo $[0, n]$, ambas fórmulas son válidas cuando 0 se sustituye por 1 o por cualquier entero positivo $< n$.

Como ejemplo de aplicación de la fórmula de sumación de Euler deduciremos la siguiente fórmula para $\log n!$.

TEOREMA 15.18. *Para cualquier entero positivo n tenemos*

$$\log n! = (n + \tfrac{1}{2}) \log n - n + C + E(n), \tag{15.49}$$

donde $0 < E(n) < 1/(8n)$ *y* $C = 1 + \int_1^\infty t^{-2}\varphi_2(t)\,dt$.

Demostración. Pongamos $f(x) = \log x$ y apliquemos el teorema 15.17 al intervalo $[1, n]$. Ello nos da

$$\sum_{k=1}^{n} \log k = \int_1^n \log x\,dx + \tfrac{1}{2}\log n + \int_1^n \frac{\varphi_2(x)}{x^2}\,dx.$$

Utilizando la relación $\int \log t\, dt = t \log t - t$ podemos poner esa ecuación en la forma

15.50 $$\log n! = (n + \tfrac{1}{2}) \log n - n + 1 + \int_1^n \frac{\varphi_2(t)}{t^2}\, dt.$$

Puesto que $|\varphi_2(t)| \leq 1/8$ la integral impropia $\int_1^\infty t^{-2}\varphi_2(t)dt$ converge absolutamente y podemos escribir

$$\int_1^n \frac{\varphi_2(t)}{t^2}\, dt = \int_1^\infty \frac{\varphi_2(t)}{t^2}\, dt - \int_n^\infty \frac{\varphi_2(t)}{t^2}\, dt.$$

Por consiguiente, la ecuación (15.50) se transforma en

$$\log n! = (n + \tfrac{1}{2}) \log n - n + C - \int_n^\infty \frac{\varphi_2(t)}{t^2}\, dt,$$

donde $C = 1 + \int_1^\infty t^{-2}\varphi_2(t)dt$. Puesto que $-1/8 < \varphi_2(t) < 0$ salvo si t es un entero o la mitad de un entero, obtenemos

$$0 < -\int_n^\infty \frac{\varphi_2(t)}{t^2}\, dt < \frac{1}{8n},$$

Esto demuestra (15.49), poniendo $E(n) = -\int_n^\infty t^{-2}\varphi_2(t)dt$.

Del teorema 15.18 podemos deducir la fórmula de Stirling que permite una estimación de $n!$.

TEOREMA 15.19. FÓRMULA DE STIRLING. *Si n es un entero positivo tenemos*

$$\sqrt{2\pi}\, n^{n+1/2}e^{-n} < n! < \sqrt{2\pi}\, n^{n+1/2}e^{-n}\left(1 + \frac{1}{4n}\right).$$

Demostración. Utilizando la ecuación (15.49) y las desigualdades para $E(n)$ obtenemos

$$\exp\left((n + \tfrac{1}{2}) \log n - n + C\right) < n! < \exp\left((n + \tfrac{1}{2}) \log n - n + C + \frac{1}{8n}\right),$$

en donde $\exp(t) = e^t$. Teniendo en cuenta que $e^x < 1 + 2x$, con $x = 1/(8n)$, podemos poner esas desigualdades en la forma

(15.51) $$An^{n+1/2}e^{-n} < n! < An^{n+1/2}e^{-n}\left(1 + \frac{1}{4n}\right),$$

donde $A = e^C$. Para terminar la demostración es preciso probar que $A = \sqrt{2\pi}$.

Deduciremos esto a partir de la desigualdad

(15.52) $$\pi n \leq \left(\frac{2^{2n}(n!)^2}{(2n)!}\right)^2 \leq \frac{\pi(2n+1)}{2},$$

descubierta por John Wallis (1616-1703). Primero demostraremos que la desigualdad de Wallis implica $A = \sqrt{2\pi}$, luego exponemos la demostración de (15.52).

Si ponemos

$$A_n = \frac{n!}{n^{n+1/2}e^{-n}}$$

la desigualdad (15.51) implica

$$A < A_n < A\left(1 + \frac{1}{4n}\right).$$

Esto prueba que $A_n \to A$ cuando $n \to \infty$. En (15.52) escribimos $n! = n^{n+1/2}e^{-n}A_n$ para conseguir

$$\pi n \leq \left(\frac{2^{2n}n^{2n+1}e^{-2n}A_n^2}{(2n)^{2n+1/2}e^{-2n}A_{2n}}\right)^2 \leq \frac{\pi(2n+1)}{2},$$

que equivale a

$$\pi \leq \frac{A_n^4}{2A_{2n}^2} \leq \pi\frac{2n+1}{2n}.$$

Haciendo que $n \to \infty$ en esta última desigualdad obtenemos

$$\pi \leq \frac{A^4}{2A^2} \leq \pi.$$

Lo que demuestra que $A^2 = 2\pi$. Puesto que $A > 0$ tenemos $A = \sqrt{2\pi}$, como se afirmó.

Queda por demostrar la desigualdad de Wallis (15.52). A tal fin introducimos los números

$$I_n = \int_0^{\pi/2} \operatorname{sen}^n t\, dt,$$

siendo n un entero no negativo cualquiera. Observemos que $I_0 = \pi/2$ e $I_1 = 1$. Para $0 \leq t \leq \pi/2$ tenemos $0 \leq \operatorname{sen} t \leq 1$; luego $0 \leq \operatorname{sen}^{n+1} t \leq \operatorname{sen}^n t$. Esto demuestra que la sucesión $\{I_n\}$ es monótona decreciente. Por tanto podemos escribir

(15.53) $$\frac{1}{I_{2n}I_{2n-1}} \leq \frac{1}{I_{2n}^2} \leq \frac{1}{I_{2n}I_{2n+1}}.$$

Calcularemos ahora cada uno de los miembros de esta desigualdad; esto nos conducirá inmediatamente a la desigualdad de Wallis.

La integración de la identidad

$$\frac{d}{dt}(\cos t \operatorname{sen}^{n+1} t) = (n+1)\operatorname{sen}^n t - (n+2)\operatorname{sen}^{n+2} t$$

en el intervalo $[0, \pi/2]$ nos da

$$0 = (n+1)I_n - (n+2)I_{n+2},$$

o

$$I_{n+2} = \frac{n+1}{n+2} I_n. \tag{15.54}$$

Utilizando esta fórmula recurrente reemplazando n por $2k - 2$ encontramos

$$\frac{I_{2k}}{I_{2k-2}} = \frac{2k-1}{2k} = \frac{2k(2k-1)}{(2k)^2}.$$

Multiplicando esas ecuaciones para $k = 1, 2, \ldots, n$ se halla

$$\prod_{k=1}^{n} \frac{I_{2k}}{I_{2k-2}} = \prod_{k=1}^{n} \frac{2k(2k-1)}{(2k)^2} = \frac{(2n)!}{2^{2n}(n!)^2}.$$

Simplificado el producto del primer miembro resulta I_{2n}/I_0. Puesto que $I_0 = \pi/2$ obtenemos

$$I_{2n} = \frac{(2n)!}{2^{2n}(n!)^2} \cdot \frac{\pi}{2}. \tag{15.55}$$

Análogamente aplicamos la fórmula de recurrencia (15.54) reemplazando n por $2k - 1$ y multiplicando las ecuaciones que resultan para $k = 1, 2, \ldots, n$ con lo que se obtiene

$$\prod_{k=1}^{n} \frac{I_{2k+1}}{I_{2k-1}} = \prod_{k=1}^{n} \frac{(2k)^2}{2k(2k+1)} = \frac{2^{2n}(n!)^2}{(2n+1)!} = \frac{1}{2n+1} \cdot \frac{\pi}{2} \cdot \frac{1}{I_{2n}}.$$

El producto del primer miembro una vez simplificado se reduce a $I_{2n+1}/I_1 = I_{2n+1}$, con lo que llegamos a

$$I_{2n} I_{2n+1} = \frac{\pi}{2(2n+1)}. \tag{15.56}$$

Puesto que $I_{2n+1} = 2n\, I_{2n-1}/(2n + 1)$, la ecuación (15.56) implica

$$I_{2n} I_{2n-1} = \frac{\pi}{4n}.$$

Utilizamos este resultado en (15.53), así como las relaciones (15.55) y (15.56). Multiplicando luego por $\pi^2/4$ obtenemos la desigualdad de Wallis (15.52).

15.23 Ejercicios

1. Si f es un polinomio de grado ≤ 2, demostrar que la fórmula de sumación de Euler (15.48) se reduce a la fórmula de los trapecios, expresada en la ecuación (15.45).
2. La constante de Euler se define mediante la fórmula

$$C = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n \right).$$

(Véase la sección 10.17 del volumen I.) Demostrar con la fórmula de sumación de Euler que

$$\sum_{k=1}^{n} \frac{1}{k} = \log n + C + \frac{1}{2n} - \frac{E(n)}{n^2},$$

donde $0 \leq E(n) \leq \frac{1}{8}$. Asimismo, demostrar que

$$C = 1 - \int_1^{\infty} \frac{t - [t]}{t^2}\, dt.$$

3) a) Si $s > 0$, $s \neq 1$, demostrar con la fórmula de sumación de Euler que

$$\sum_{k=1}^{n} \frac{1}{k^s} = \frac{n^{1-s}}{1 - s} + C(s) + s \int_n^{\infty} \frac{t - [t]}{t^{s+1}}\, dt,$$

siendo

$$C(s) = 1 + \frac{1}{s - 1} - s \int_1^{\infty} \frac{t - [t]}{t^{s+1}}\, dt.$$

b) Si $s > 1$, demostrar que $C(s) = \zeta(s)$, siendo ζ la función zeta de Riemann definida para $s > 1$ mediante la serie

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}.$$

Esta serie diverge para $s \leq 1$. No obstante, puesto que la fórmula que da $C(s)$ en la parte a) tiene sentido para $0 < s < 1$, puede utilizarse para extender la definición de $\zeta(s)$ al intervalo abierto $0 < s < 1$. Así pues, para $s > 0$ y $s \neq 1$ tenemos la fórmula

$$\zeta(s) = 1 + \frac{1}{s-1} - s\int_1^\infty \frac{t-[t]}{t^{s+1}}\,dt.$$

Esto es un *teorema* si $s > 1$, y una *definición* si $0 < s < 1$.

En los ejercicios del 4 al 6, φ_2 es la función introducida en 15.22.

4. a) Emplear la fórmula de sumación de Euler para demostrar que

$$\sum_{k=1}^{n} \log^2 k = (n + \tfrac{1}{2})\log^2 n - 2n\log n + 2n - 2 + 2\int_1^n \varphi_2(x)\,\frac{\log x - 1}{x^2}\,dx.$$

b) Utilizar la parte a) para deducir que para $n > e$ tenemos

$$\sum_{k=1}^{n} \log^2 k = (n + \tfrac{1}{2})\log^2 n - 2n\log n + 2n + A - E(n),$$

siendo A una constante y $0 < E(n) < E(n) < \dfrac{\log n}{4n}$.

5. a) Con la forma de sumación de Euler demostrar que

$$\sum_{k=1}^{n} \frac{\log k}{k} = \frac{1}{2}\log^2 n + \frac{1}{2}\frac{\log n}{n} - \int_1^n \frac{2\log x - 3}{x^3}\,\varphi_2(x)\,dx.$$

b) Utilizar la parte a) para deducir que para que $n > e^{3/2}$ tenemos

$$\sum_{k=1}^{n} \frac{\log k}{k} = \frac{1}{2}\log^2 n + \frac{1}{2}\frac{\log n}{n} + A - E(n),$$

siendo A una constante y $0 < E(n) < \dfrac{\log n}{8n^2}$.

6. a) Si $n > 2$ utilizar la fórmula de sumación de Euler para demostrar que

$$\sum_{k=2}^{n} \frac{1}{k\log k}$$

$$= \log(\log n) + \frac{1}{2n\log n} + \frac{1}{4\log 2} - \log(\log 2) - \int_2^n \varphi_2(x)\,\frac{2 + 3\log x + 2\log^2 x}{(x\log x)^3}\,dx.$$

b) Utilizar la parte a) para deducir que para $n > 2$ tenemos

$$\sum_{k=2}^{n} \frac{1}{k \log k} = \log(\log n) + A + \frac{1}{2n \log n} - E(n),$$

siendo A una constante y $0 < E(n) < \dfrac{1}{4n^2 \log n}$.

7. a) Si $a > 0$ y $p > 0$, demostrar con la fórmula de sumación de Euler que

$$\sum_{k=0}^{\infty} e^{-ak^p} = \frac{\Gamma\left(1 + \frac{1}{p}\right)}{a^{1/p}} + \frac{1}{2} - ap \int_0^{\infty} \varphi_1(x) x^{p-1} e^{-ax^p}\, dx,$$

siendo Γ la función gamma.

b) Mediante la parte a) demostrar que

$$\sum_{k=0}^{\infty} e^{-ak^p} = \frac{\Gamma\left(1 + \frac{1}{p}\right)}{a^{1/p}} + \theta, \qquad \text{donde} \quad 0 < \theta < 1.$$

8. Deducir los límites siguientes con la ayuda de la fórmula de Stirling o de la de Wallis o de ambas.

a) $\lim\limits_{n \to \infty} \dfrac{n}{(n!)^{1/n}} = e.$

b) $\lim\limits_{n \to \infty} \dfrac{(n!)^2 2^{2n}}{(2n)!\, \sqrt{n}} = \sqrt{\pi}.$

c) $\lim\limits_{n \to \infty} (-1)^n \dbinom{-\frac{1}{2}}{n} n = \dfrac{1}{\sqrt{\pi}}.$

9. Sea $I_n = \int_0^{\pi/2} \operatorname{sen}^n t\, dt$, siendo n un entero no negativo. En la sección 15.22 se demostró que la sucesión $\{I_n\}$ satisface la fórmula de recurrencia

$$I_{n+2} = \frac{n+1}{n+2} I_n.$$

Pongamos $f(n) = \frac{1}{2}\sqrt{\pi}\, \Gamma\left(\frac{n+1}{2}\right) \Big/ \Gamma\left(\frac{n}{2} + 1\right)$, siendo Γ la función gamma.

a) Utilizar la ecuación funcional $\Gamma(s+1) = s\Gamma(s)$ para demostrar que

$$f(n+2) = \frac{n+1}{n+2} f(n).$$

b) Con la parte a) deducir que

$$\int_0^{\pi/2} \operatorname{sen}^n t\, dt = \frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)}.$$

REFERENCIAS CITADAS

Esta pequeña lista contiene tan sólo unos pocos libros para estudiar más ampliamente los principios generales del análisis numérico. Todos ellos contienen a su vez otras referencias y citas de trabajos más especializados. Recomendamos especialmente la lista de tablas dada en el libro que se indica en la referencia 9.

Libros

1. A. D. Booth, *Numerical Methods,* Academic Press, Nueva York, 1958; 3.ª ed., Plenum Press, Nueva York, 1966.
2. P. J. Davis, *Interpolation and Approximation,* Blaisdell, Waltham, Mass., 1963.
3. D. R. Hartree, *Numerical Analysis,* Oxford Univ. Press (Clarendon), Londres y Nueva York, 1958.
4. F. B. Hildebrand, *Introduction to Numerical Analysis,* McGraw-Hill, Nueva York, 1956.
5. A. S. Householder, *Principles of Numerical Analysis,* McGraw-Hill, Nueva York, 1953.
6. W. E. Milne, *Numerical Calculus,* Princeton Univ. Press, Princeton, N.J., 1950.
7. J. B. Scarborough, *Numerical Mathematical Analysis,* Johns Hopkins Press, Baltimore, Md., 1958, 6.ª ed., 1966.
8. J. Todd, *Introduction to the Constructive Theory of Functions,* Academic Press, Nueva York, 1963.
9. J. Todd(ed.), *Survey of Numerical Analysis,* McGraw-Hill, Nueva York, 1962.

Tablas

10. L. J. Comrie (ed.), *Chambers' Six-figure Mathematical Tables,* W. & R. Chambers, Londres y Edimburgo, 1949.
11. L. J. Comrie, «Interpolation and Allied Tables», 2.ª reimpresión rev. del *Nautical Almanac* para 1937, H. M. Stationery Office, Londres, 1948.
12. A. J. Fletcher, J. C. P. Miller, y L. Rosenhead, *Index of Mathematical Tables,* McGraw-Hill, Nueva York, 1946.
13. *Tables of Lagrangian Interpolation Coefficients,* Natl. Standards Columbia Press Series, Vol. 4, Columbia Univ. Press, Nueva York, 1944.

SOLUCIONES A LOS EJERCICIOS

Capítulo 1

1.5 Ejercicios (pág. 8)

1. Si
2. Si
3. Si
4. Si
5. No
6. Si
7. Si
8. Si
9. Si
10. Si
11. No
12. Si
13. Si
14. No
15. Si
16. Si
17. Si
18. Si
19. Si
20. Si
21. Si
22. Si
23. No
24. Si
25. No
26. Si
27. Si
28. Si

31. a) No b) No c) No d) No

1.10 Ejercicios (pág. 16)

1. Si ; 2
2. Si ; 2
3. Si ; 2
4. Si ; 2
5. Si ; 1
6. No
7. No
8. No
9. Si ; 1
10. Si ; 1
11. Si ; n
12. Si ; n
13. Si ; n
14. Si ; n
15. Si ; n
16. Si ; n
17. Si ; dim $= 1 + \frac{1}{2}n$ si n es par, $\frac{1}{2}(n+1)$ si n es impar
18. Si ; dim $= \frac{1}{2}n$ si n es par , $\frac{1}{2}(n+1)$ si n es impar
19. Si ; $k + 1$
20. No
21. a) dim $= 3$ b) dim $= 3$ c) dim $= 2$ d) dim $= 2$
23. a) si $a \neq 0$ y $b \neq 0$, el conjunto es independiente, dim $= 3$; si a ó b es cero, el conjunto es dependiente; dim $= 2$ (b) independiente , dim $= 2$ c) si $a \neq 0$, independiente , dim $= 3$; si $a = 0$, dependiente, dim $= 2$ d) independiente; dim $= 3$ e) dependiente; dim $= 2$ f) independiente dim $= 2$ g) independiente dim $= 2$ h) dependiente; dim $= 2$ i) independiente; dim $= 2$ j) independiente; dim $= 2$

1.13 Ejercicios (pág. 24)

1. a) No b) No c) No d) No e) Si
8. a) $\frac{1}{2}\sqrt{e^2+1}$ b) $g(x) = b\left(x - \frac{e^2+1}{4}\right)$, b arbitrario
10. b) $\frac{(n+1)(2n+1)}{6n}a + \frac{n+1}{2}b$ c) $g(t) = a\left(t - \frac{2n+1}{3n}\right)$, a arbitrario
11. c) 43 d) $g(t) = a(1 - \frac{2}{3}t)$, a arbitrario
12. a) No b) No c) No d) No
13. c) 1 d) $e^2 - 1$
14. c) $n!/2^{n+1}$

1.17 Ejercicios (pág. 36)

1. a) y b) $\frac{1}{3}\sqrt{3}\,(1, 1, 1)$, $\frac{1}{6}\sqrt{6}\,(1, -2, 1)$

2. a) $\frac{1}{2}\sqrt{2}\,(1, 1, 0, 0)$, $\frac{1}{6}\sqrt{6}\,(-1, 1, 2, 0)$, $\frac{1}{6}\sqrt{3}\,(1, -1, 1, 3)$

   b) $\frac{1}{3}\sqrt{3}\,(1, 1, 0, 1)$, $\dfrac{1}{\sqrt{42}}(1, -2, 6, 1)$

6. $\frac{2}{3} - \frac{1}{2}\log^2 3$
7. $e^2 - 1$
8. $\frac{1}{2}(e - e^{-1}) + \dfrac{3}{e}x$; $1 - 7e^{-2}$
9. $\pi - 2\,\text{sen}\,x$
10. $\frac{3}{4} - \frac{1}{4}x$

Capítulo 2

2.4 Ejercicios (pág. 44)

1. Lineal ; dimensión del núcleo 0, rango 2
2. Lineal ; dimensión del núcleo 0, rango 2
3. Lineal ; dimensión del núcleo 1, rango 1
4. Lineal ; dimensión del núcleo 1, rango 1
5. No lineal
6. No lineal
7. No lineal
8. No lineal
9. Lineal ; dimensión del núcleo 0, rango 2
10. Lineal ; dimensión del núcleo 0, rango 2
11. Lineal ; dimensión del núcleo 0, rango 2
12. Lineal ; dimensión del núcleo 0, rango 2
13. No lineal
14. Lineal ; dimensión del núcleo 0, rango 2
15. No lineal
16. Lineal ; dimensión del núcleo 0, rango 3
17. Lineal ; dimensión del núcleo 1, rango 2
18. Lineal ; dimensión del núcleo 0, rango 3
19. No lineal
20. No lineal
21. No lineal
22. No lineal
23. Lineal ; dimensión del núcleo 1, rango 2
24. Lineal ; dimensión del núcleo 0, rango $n + 1$
25. Lineal ; dimensión del núcleo 1, rango infinito
26. Lineal ; dimensión del núcleo infinita, rango 2
27. Lineal ; dimensión del núcleo 2, rango infinito
28. $N(T)$ es el conjunto de las sucesiones constantes; $T(V)$ es el conjunto de las sucesiones con límite 0
29. d) $\{1, \cos x, \text{sen}\,x\}$ es una base para $T(V)$; $\dim T(V) = 3$ e) $N(T) = S$ f) Si $T(f) = cf$ siendo $c \neq 0$, entonces $c \in T(V)$ con lo que tenemos $f(x) = c_1 + c_2 \cos x + c_3\,\text{sen}\,x$; si $c_1 = 0$, entonces $c = \pi$ y $f(x) = c_1 \cos x + c_2\,\text{sen}\,x$, donde c_1, c_2 son arbitrarias y no simultáneamente nulas; si $c_1 \neq 0$, entonces $c = 2\pi$ y $f(x) = c_1$, donde c_1 no es nula pero arbitraria.

2.8 Ejercicios (pág. 53)

3. Si ; $x = v, \quad y = u$
4. Si ; $x = u, \quad y = -v$
5. No
6. No
7. No
8. Si ; $x = \log u, \quad y = \log v$
9. No
10. Si ; $x = u - 1, \quad y = v - 1$
11. Si ; $x = \frac{1}{2}(v + u), \quad y = \frac{1}{2}(v - u)$
12. Si ; $x = \frac{1}{3}(v + u), \quad y = \frac{1}{3}(2v - u)$
13. Si ; $x = w, \quad y = v, \quad z = u$
14. No
15. Si ; $x = u, \quad y = \frac{1}{2}v, \quad z = \frac{1}{3}w$
16. Si ; $x = u, \quad y = v, \quad z = w - u - v$
17. Si ; $x = u - 1, \quad y = v - 1, \quad z = w + 1$
18. Si ; $x = u - 1, \quad y = v - 2, \quad z = w - 3$
19. Si ; $x = u, \quad y = v - u, \quad z = w - v$
20. Si ; $x = \frac{1}{2}(u - v + w), \quad y = \frac{1}{2}(v - w + u); \quad z = \frac{1}{2}(w - u + v)$
25. $(S + T)^2 = S^2 + ST + TS + T^2$;
$(S + T)^3 = S^3 + TS^2 + STS + S^2T + ST^2 + TST + T^2S + T^3$
26. a) $(ST)(x, y, z) = (x + y + z, x + y, x)$; $(TS)(x, y, z) = (z, z + y, z + y + x)$;
$(ST - TS)(x, y, z) = (x + y, x - z, -y - z)$; $S^2(x, y, z) = (x, y, z)$;
$T^2(x, y, z) = (x, 2x + y, 3x + 2y + z)$;
$(ST)^2(x, y, z) = (3x + 2y + z, 2x + 2y + z, x + y + z)$;
$(TS)^2(x, y, z) = (x + y + z, x + 2y + 2z, x + 2y + 3z)$;
$(ST - TS)^2 = (2x + y - z, x + 2y + z, -x + y + 2z)$;
b) $S^{-1}(u, v, w) = (w, v, u)$; $T^{-1}(u, v, w) = (u, v - u, w - v)$;
$(ST)^{-1}(u, v, w) = (w, v - w, u - v)$; $(TS)^{-1}(u, v, w) = (w - v, v - u, u)$
c) $(T - I)(x, y, z) = (0, x, x + y)$; $(T - I)^2(x, y, z) = (0, 0, x)$;
$(T - I)^n(x, y, z) = (0, 0, 0)$ si $n \geq 3$
28. a) $Dp(x) = 3 - 2x + 12x^2$; $Tp(x) = 3x - 2x^2 + 12x^3$; $(DT)p(x) = 3 - 4x + 36x^2$;
$(TD)p(x) = -2x + 24x^2$; $(DT - TD)p(x) = 3 - 2x + 12x^2$;
$(T^2D^2 - D^2T^2)p(x) = 8 - 192x$ b) $p(x) = ax$, a un escalar cualquiera
c) $p(x) = ax^2 + b$, a y b escalares cualesquiera d) Todo p de V
31. a) $Rp(x) = 2$; $Sp(x) = 3 - x + x^2$; $Tp(x) = 2x + 3x^2 - x^3 + x^4$;
$(ST)p(x) = 2 + 3x - x^2 + x^3$; $(TS)p(x) = 3x - x^2 + x^3$; $(TS)^2p(x) = 3x - x^2 + x^3$;
$(T^2S^2)p(x) = -x^2 + x^3$; $(S^2T^2)p(x) = 2 + 3x - x^2 + x^3$; $(TRS)p(x) = 3x$;
$(RST)p(x) = 2$
b) $N(R) = \{p \mid p(0) = 0\}$; $R(V) = \{p \mid p$ es constante$\}$; $N(S) = \{p \mid p$ es constante $\}$;
$S(V) = V$; $N(T) = \{O\}$; $T(V) = \{p \mid p(0) = 0\}$ c) $T^{-1} = S$
d) $(TS)^n = I - R$; $S^nT^n = I$
32. T no es uno a uno en V porque aplica toda sucesión constante en la misma sucesión.

2.12 Ejercicios (pág. 62)

1. a) La matriz identidad $I = (\delta_{jk})$, siendo $\delta_{jk} = 1$ si $j = k$, y $\delta_{jk} = 0$ si $j \neq k$
   b) La matriz cero $O = (a_{jk})$, en la que cada elemento $a_{jk} = 0$.
   c) La matriz $(c\delta_{jk})$, siendo (δ_{jk}) la matriz identidad de la parte a)

2. a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ c) $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

3. a) $-5\boldsymbol{i} + 7\boldsymbol{j},\quad 9\boldsymbol{i} - 12\boldsymbol{j}$
   b) $\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ c) $\begin{bmatrix} -\frac{7}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{7}{4} \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$

4. $\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$

5. a) $3\boldsymbol{i} + 4\boldsymbol{j} + 4\boldsymbol{k}$; dimensión del núcleo 0, rango 3 b) $\begin{bmatrix} -1 & -1 & 2 \\ 1 & -3 & 3 \\ -1 & -5 & 5 \end{bmatrix}$

6. $\begin{bmatrix} 2 & 0 & -2 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$

7. a) $T(4\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}) = (0, -2)$; dimensión del núcleo 1, rango 2 b) $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$
   c) $\begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix}$ d) $e_1 = \boldsymbol{j},\quad e_2 = \boldsymbol{k},\quad e_3 = \boldsymbol{i},\quad w_1 = (1, 1),\quad w_2 = (1, -1)$

8. a) $(5, 0, -1)$; dimensión del núcleo 0, rango 2 b) $\begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$
   c) $e_1 = \boldsymbol{i},\quad e_2 = \boldsymbol{i} + \boldsymbol{j},\quad w_1 = (1, 0, 1),\quad w_2 = (0, 0, 2),\quad w_3 = (0, 1, 0)$

9. a) $(-1, -3, -1)$; dimensión del núcleo 0, rango 2 b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$
   c) $e_1 = \boldsymbol{i},\quad e_2 = \boldsymbol{j} - \boldsymbol{i},\quad w_1 = (1, 0, 1),\quad w_2 = (0, 1, 0),\quad w_3 = (0, 0, 1)$

10. a) $e_1 - e_2$; dimensión del núcleo 0, rango 2 b) $\begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$ c) $a = 5,\quad b = 4$

11. $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

12. $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

16. $\begin{bmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 & 0 & -2 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

17. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$

18. $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$, $\begin{bmatrix} -5 & -12 \\ 12 & -5 \end{bmatrix}$

19. a) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

d) $\begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ e) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$ f) $\begin{bmatrix} 0 & 0 & -8 & 0 \\ 0 & 0 & 0 & -48 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

20. Elíjase $(x^3, x^2, x, 1)$ como base para V, y (x^2, x) como base para W. Entonces la matriz de TD es

$$\begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}$$

2.16 Ejercicios (pág. 70)

1. $B + C = \begin{bmatrix} 3 & 4 \\ 0 & 2 \\ 6 & -5 \end{bmatrix}$, $AB = \begin{bmatrix} 15 & -14 \\ -15 & 14 \end{bmatrix}$, $BA = \begin{bmatrix} -1 & 4 & -2 \\ -4 & 16 & -8 \\ 7 & -28 & 14 \end{bmatrix}$,

$AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $CA = \begin{bmatrix} 0 & 0 & 0 \\ 2 & -8 & 4 \\ 4 & -16 & 8 \end{bmatrix}$, $A(2B - 3C) = \begin{bmatrix} 30 & -28 \\ -30 & 28 \end{bmatrix}$

2. a) $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$, a y b arbitrarios b) $\begin{bmatrix} -2a & a \\ -2b & b \end{bmatrix}$, a y b arbitrarios

3. a) $a = 9$, $b = 6$, $c = 1$, $d = 5$ b) $a = 1$, $b = 6$, $c = 0$, $d = -2$

4. a) $\begin{bmatrix} -9 & -2 & -10 \\ 6 & 14 & 8 \\ -7 & 5 & -5 \end{bmatrix}$ b) $\begin{bmatrix} -3 & 5 & -4 \\ 0 & 3 & 24 \\ 12 & -27 & 0 \end{bmatrix}$

6. $A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$

7. $A^n = \begin{bmatrix} \cos n\theta & -\operatorname{sen} n\theta \\ \operatorname{sen} n\theta & \cos n\theta \end{bmatrix}$.

8. $A^n = \begin{bmatrix} 1 & n & \dfrac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 \\ -100 & 1 \end{bmatrix}$

10. $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, siendo b y c arbitrarios y a es una solución cualquiera de la ecuación $a^2 = -bc$

11. b) $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, siendo a arbitrario

12. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, y $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, siendo b y c arbitrarios y a es una solución cualquiera de la ecuación $a^2 = 1 - bc$

13. $C = \begin{bmatrix} \frac{15}{2} & \frac{13}{2} \\ 8 & 7 \end{bmatrix}$, $D = \begin{bmatrix} \frac{33}{4} & \frac{19}{4} \\ \frac{43}{4} & \frac{25}{4} \end{bmatrix}$

14. b) $(A + B)^2 = A^2 + AB + BA + B^2$; $(A + B)(A - B) = A^2 + BA - AB - B^2$;
c) Para aquellos que conmutan

2.20 Ejercicios (pág. 83)

1. $(x, y, z) = (\frac{8}{5}, -\frac{7}{5}, \frac{8}{5})$
2. Ninguna solución
3. $(x, y, z) = (1, -1, 0) + t(-3, 4, 1)$
4. $(x, y, z) = (1, -1, 0) + t(-3, 4, 1)$
5. $(x, y, z, u) = (1, 1, 0, 0) + t(1, 14, 5, 0)$
6. $(x, y, z, u) = (1, 8, 0, -4) + t(2, 7, 3, 0)$
7. $(x, y, z, u, v) = t_1(-1, 1, 0, 0, 0) + t_2(-1, 0, 3, -3, 1)$
8. $(x, y, z, u) = (1, 1, 1, -1) + t_1(-1, 3, 7, 0) + t_2(4, 9, 0, 7)$
9. $(x, y, z) = (\frac{4}{3}, \frac{2}{3}, 0) + t(5, 1, -3)$
10. a) $(x, y, z, u) = (1, 6, 3, 0) + t_1(4, 11, 7, 0) + t_2(0, 0, 0, 1)$
b) $(x, y, z, u) = (\frac{3}{11}, 4, \frac{19}{11}, 0) + t(4, -11, 7, 22)$

12. $\begin{bmatrix} -1 & 2 & 1 \\ 5 & -8 & -6 \\ -3 & 5 & 4 \end{bmatrix}$

14. $\begin{bmatrix} 14 & 8 & 3 \\ 8 & 5 & 2 \\ 3 & 2 & 1 \end{bmatrix}$

13. $\begin{bmatrix} -\frac{5}{3} & \frac{2}{3} & \frac{4}{3} \\ -1 & 0 & 1 \\ \frac{7}{3} & -\frac{1}{3} & -\frac{5}{3} \end{bmatrix}$

15. $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

16. $\begin{bmatrix} 0 & \frac{1}{2} & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ -3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 9 & 0 & -3 & 0 & 1 & 0 \end{bmatrix}$

2.21 Ejercicios varios sobre matrices (pág. 84)

3. $P = \begin{bmatrix} 2 & 1 \\ 5 & -1 \end{bmatrix}$

4. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, y $\begin{bmatrix} a & b \\ c & 1 - a \end{bmatrix}$, donde b y c son arbitrarios y a es una solución cualquiera de la ecuación cuadrática $a^2 - a + bc = 0$.

10. a) $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}$, $\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$,

$\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$, $\begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$.

Capítulo 3

3.6 Ejercicios (pág. 97)

1. a) 6 b) 76 c) $a^3 - 4a$
2. a) 1 b) 1 c) 1
3. b) $(b-a)(c-a)(c-b)(a+b+c)$ y $(b-a)(c-a)(c-b)(ab+ac+bc)$
4. a) 8 (b) $(b-a)(c-a)(d-a)(c-b)(d-b)(d-c)$
   c) $(b-a)(c-a)(d-a)(c-b)(d-b)(d-c)(a+b+c+d)$
   d) $a(a^2-4)(a^2-16)$ e) -160

7. $F' = \begin{vmatrix} f_1' & f_2' & f_3' \\ g_1 & g_2 & g_3 \\ h_1 & h_2 & h_3 \end{vmatrix} + \begin{vmatrix} f_1 & f_2 & f_3 \\ g_1' & g_2' & g_3' \\ h_1 & h_2 & h_3 \end{vmatrix} + \begin{vmatrix} f_1 & f_2 & f_3 \\ g_1 & g_2 & g_3 \\ h_1' & h_2' & h_3' \end{vmatrix}$

8. b) Si $F = \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{vmatrix}$ entonces $F' = \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1''' & f_2''' & f_3''' \end{vmatrix}$

10 $\det A = 16$, $\det(A^{-1}) = \dfrac{1}{16}$, $A^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{3}{4} & \frac{1}{8} & \frac{1}{16} \\ 0 & \frac{1}{2} & -\frac{3}{4} & \frac{1}{8} \\ 0 & 0 & \frac{1}{2} & -\frac{3}{4} \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$

3.11 Ejercicios (pág. 104)

6. $\det A = (\det B)(\det D)$
7. a) Independientes b) Independientes c) Dependientes

3.17 Ejercicios (pág. 116)

1. a) $\begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 2 & -1 & 1 \\ -6 & 3 & 5 \\ -4 & -2 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 109 & 113 & -41 & -13 \\ -40 & -92 & 74 & 16 \\ -41 & -79 & 7 & 47 \\ -50 & 38 & 16 & 20 \end{bmatrix}$

2. a) $-\frac{1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$ b) $\frac{1}{8}\begin{bmatrix} 2 & -6 & -4 \\ -1 & 3 & -2 \\ 1 & 5 & 2 \end{bmatrix}$ c) $\frac{1}{306}\begin{bmatrix} 109 & -40 & -41 & -50 \\ 113 & -92 & -79 & 38 \\ -41 & 74 & 7 & 16 \\ -13 & 16 & 47 & 20 \end{bmatrix}$

3. a) $\lambda = 2, \quad \lambda = -3$ b) $\lambda = 0, \quad \lambda = \pm 3$ c) $\lambda = 3, \quad \lambda = \pm i$

5. a) $x = 0, \quad y = 1, \quad z = 2$ b) $x = 1, \quad y = 1, \quad z = -1$

6. b) $\det\begin{bmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{bmatrix} = 0; \quad \det\begin{bmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{bmatrix} = 0$

c) $\det\begin{bmatrix} (x - x_1)^2 + (y - y_1)^2 & (x - x_1) & (y - y_1) \\ (x_2 - x_1)^2 + (y_2 - y_1)^2 & (x_2 - x_1) & (y_2 - y_1) \\ (x_3 - x_1)^2 + (y_3 - y_1)^2 & (x_3 - x_1) & (y_3 - y_1) \end{bmatrix} = 0;$

$$\det\begin{bmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{bmatrix} = 0$$

Capítulo 4

4.4 Ejercicios (pág. 125)

5. Autofunciones : $f(t) = Ct^{\lambda}$, donde $C \neq 0$
6. El polinomio constante no cero
7. Autofunciones : $f(t) = Ce^{t/\lambda}$, donde $C \neq 0$
8. Autofunciones : $f(t) = Ce^{\frac{1}{2}t^2/\lambda}$, donde $C \neq 0$
0. Los autovectores pertenecientes a $\lambda = 0$ son todos sucesiones constantes con límite $a \neq 0$.

Los autovectores pertenecientes a $\lambda = -1$ son todos sucesiones no constantes con límite $a = 0$

4.8 Ejercicios (pág. 132)

		Autovalor	*Autovectores*	dim $E(\lambda)$
1.	(a)	1, 1	$(a, b) \neq (0, 0)$	2
	(b)	1, 1	$t(1, 0), t \neq 0$	1
	(c)	1, 1	$t(0, 1), t \neq 0$	1
	(d)	2	$t(1, 1)\ \ t \neq 0$	1
		0	$t(1, -1), t \neq 0$	1
2.		$1 + \sqrt{ab}$	$t(\sqrt{a}, \sqrt{b}), t \neq 0$	1
		$1 - \sqrt{ab}$	$t(\sqrt{a}, -\sqrt{b}), t \neq 0$	1

3. Si el campo de escalares es el conjunto de los números reales **R**, entonces solamente existen autovectores reales cuando $\operatorname{sen}\theta = 0$, en este caso habrá dos autovectores iguales, $\lambda_1 = \lambda_2 = \cos\theta$, donde $\cos\theta = 1$ o -1
En este caso todo vector no cero es un autovector $\dim E(\lambda_1) = \dim E(\lambda_2) = 2$.
Si el campo de escalares es el conjunto de los números complejos **C**, entonces los autovectores son $\lambda_1 = \cos\theta + i\operatorname{sen}\theta, \lambda_2 = \cos\theta - i\operatorname{sen}\theta$. Si $\operatorname{sen}\theta = 0$, son reales e iguales. Si $\operatorname{sen}\theta \neq 0$ son complejos conjugados distintos; los autovectores pertenecientes a λ_1 son $t(i, 1)$, $t \neq 0$; los pertenecientes a λ_2 son $t(1, i)$, $t \neq 0$; $\dim E(\lambda_1) = \dim E(\lambda_2) = 1$.

4. $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, donde b y c son arbitrarios y a es una de las soluciones de la ecuación $a^2 = 1 - bc$.

5. Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, y sea $\Delta = (a - d)^2 + 4bc$. Los autovectores son reales y distintos si $\Delta > 0$, reales e iguales si $\Delta = 0$, complejos conjugados si $\Delta < 0$.

6. $a = b = c = d = e = f = 1$.

		Autovalores	*Autovectores*	dim $E(\lambda)$
7.	(a)	1, 1, 1	$t(0, 0, 1), t \neq 0$	1
	(b)	1	$t(1, -1, 0), t \neq 0$	1
		2	$t(3, 3, -1), t \neq 0$	1
		21	$t(1, 1, 6), t \neq 0$	1
	(c)	1	$t(3, -1, 3), t \neq 0$	1
		2, 2	$t(2, 2, -1), t \neq 0$	1

8. 1, 1, -1, -1 respectivamente para cada matriz

4.10 Ejercicios (pág. 139)

2. a) Autovalores 1, 3; $C = \begin{bmatrix} -2c & 0 \\ c & d \end{bmatrix}$, donde $cd \neq 0$

(b) Autovalores 6, −1; $C = \begin{bmatrix} 2a & b \\ 5a & -b \end{bmatrix}$, donde $ab \neq 0$

c) Autovalores 3, 3; si existe el no singular C, entonces $C^{-1}AC = 3I$, siendo $AC = 3C$, $A = 3I$

d) Autovalores 1, 1; si existe el no singular C, entonces $C^{-1}AC = I$, siendo $AC = C$, $A = I$

3. $C = A^{-1}B$.

4. a) Autovalores 1, 1, −1; autovectores (1, 0, 1), (0, 1, 0), (1, 0, −1);

$$C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

b) Autovalores 2, 2, 1; autovectores (1, 0, −1), (0, 1, −1), (1, −1, 1);

$$C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

5. a) Autovalores 2, 2; autovectores $t(1, 0)$, $t \neq 0$. Si $C = \begin{bmatrix} a & b \\ -b & 0 \end{bmatrix}$, $b \neq 0$,

entonces $C^{-1}AC = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$

b) Autovalores 3, 3; autovectores $t(1, 1)$, $t \neq 0$. Si $C = \begin{bmatrix} a & b \\ a+b & b \end{bmatrix}$, $b \neq 0$,

entonces $C^{-1}AC = \begin{bmatrix} 3 & 0 \\ 1 & 3 \end{bmatrix}$

6. Autovalores 1, 1, 1; autovectores $t(1, -1, -1)$, $t \neq 0$

Capítulo 5

5.5 Ejercicios (pág. 146)

3. b) T^n es hermitiana si n es par, si n es impar es anti-hermitiana
7. a) Simétrica b) Ni lo uno ni lo otro c) Simétrica d) Simétrica
9. d) $Q(x + ty) = Q(x) + t\bar{t}Q(y) + \bar{t}(T(x), y) + t(T(y), x)$

5.11 Ejercicios (pág. 154)

1. a) Simétrica hermitiana

b) Ninguno de los cuatro tipos
c) Anti-simétrica
d) Anti-simétrica y anti-hermitiana

4. b) $\begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$

5. Autovalores $\lambda_1 = 0$, $\lambda_2 = 25$; autovectores ortonormales $\boldsymbol{u}_1 = \frac{1}{5}(4, -3)$, $\boldsymbol{u}_2 = \frac{1}{5}(3, 4)$.

$$C = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$$

6. Autovalores $\lambda_1 = 2i$, $\lambda_2 = -2i$; autovectores ortonormales

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}(1, -i), \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{2}}(1, i). \qquad C = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$$

7. Autovalores $\lambda_1 = 1$, $\lambda_2 = 3$, $\lambda_3 = -4$; autovectores ortonormales

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{10}}(1, 0, 3), \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{14}}(3, 2, -1), \quad \boldsymbol{u}_3 = \frac{1}{\sqrt{35}}(3, -5, -1).$$

$$C = \begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{14}} & \frac{3}{\sqrt{35}} \\ 0 & \frac{2}{\sqrt{14}} & \frac{-5}{\sqrt{35}} \\ \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{14}} & \frac{-1}{\sqrt{35}} \end{bmatrix}$$

8. Autovalores $\lambda_1 = 1$, $\lambda_2 = 6$, $\lambda_3 = -4$; autovectores ortonormales

$$\boldsymbol{u}_1 = \tfrac{1}{5}(0, 4, -3), \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{50}}(5, 3, 4), \quad \boldsymbol{u}_3 = \frac{1}{\sqrt{50}}(5, -3, -4).$$

$$C = \frac{1}{\sqrt{50}}\begin{bmatrix} 0 & 5 & 5 \\ 4\sqrt{2} & 3 & -3 \\ -3\sqrt{2} & 4 & -4 \end{bmatrix}$$

9. a), b), c) son unitarios b), c) son ortogonales
11. a) Autovalores $\lambda_1 = ia$, $\lambda_2 = -ia$; autovectores ortonormales

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}(1, i), \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{2}}(1, -i). \qquad \text{b) } C = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

5.15 Ejercicios (pág. 166)

1. a) $A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$ b) $\lambda_1 = 0, \quad \lambda_2 = 5$ c) $u_1 = \frac{1}{\sqrt{5}}(1, -2), \quad u_2 = \frac{1}{\sqrt{5}}(2, 1)$

d) $C = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$

2. a) $A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$ b) $\lambda_1 = \frac{1}{2}, \quad \lambda_2 = -\frac{1}{2}$

c) $u_1 = \frac{1}{\sqrt{2}}(1, 1), \quad u_2 = \frac{1}{\sqrt{2}}(1, -1)$ d) $C = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

3. a) $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ b) $\lambda_1 = \sqrt{2}, \quad \lambda_2 = -\sqrt{2}$

c) $u_1 = t(1 + \sqrt{2}, 1), \quad u_2 = t(-1, 1 + \sqrt{2})$, donde $t = 1/\sqrt{4 + 2\sqrt{2}}$

d) $C = t\begin{bmatrix} 1 + \sqrt{2} & -1 \\ 1 & 1 + \sqrt{2} \end{bmatrix}$, donde $t = 1/\sqrt{4 + 2\sqrt{2}}$

4. a) $A = \begin{bmatrix} 34 & -12 \\ -12 & 41 \end{bmatrix}$ b) $\lambda_1 = 50, \quad \lambda_2 = 25$

c) $u_1 = \frac{1}{5}(3, -4), \quad u_2 = \frac{1}{5}(4, 3)$ d) $C = \frac{1}{5}\begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$

5. a) $A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$ b) $\lambda_1 = 0, \quad \lambda_2 = \frac{3}{2}, \quad \lambda_3 = -\frac{1}{2}$

c) $u_1 = \frac{1}{\sqrt{3}}(1, -1, -1), \; u_2 = \frac{1}{\sqrt{6}}(2, 1, 1), \quad u_3 = \frac{1}{\sqrt{2}}(0, 1, -1)$

c) $C = \frac{1}{\sqrt{6}}\begin{bmatrix} \sqrt{2} & 2 & 0 \\ -\sqrt{2} & 1 & \sqrt{3} \\ -\sqrt{2} & 1 & -\sqrt{3} \end{bmatrix}$

6. a) $A = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix}$ b) $\lambda_1 = 1, \quad \lambda_2 = 3, \quad \lambda_3 = -2$

c) $u_1 = (0, 1, 0)$, $u_2 = \frac{1}{\sqrt{5}}(2, 0, 1)$, $u_3 = \frac{1}{\sqrt{5}}(1, 0, -2)$

d) $C = \frac{1}{\sqrt{5}}\begin{bmatrix} 0 & 2 & 1 \\ \sqrt{5} & 0 & 0 \\ 0 & 1 & -2 \end{bmatrix}$

7. a) $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ b) $\lambda_1 = \lambda_2 = -1$, $\lambda_3 = 8$

c) $u_1 = \frac{1}{\sqrt{2}}(1, 0, -1)$, $u_2 = \frac{1}{3\sqrt{2}}(-1, 4, -1)$, $u_3 = \frac{1}{3}(2, 1, 2)$

d) $C = \frac{1}{3\sqrt{2}}\begin{bmatrix} 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \\ -3 & -1 & 2\sqrt{2} \end{bmatrix}$

8. Elipse; centro en $(0, 0)$
9. Hipérbola; centro en $(-\frac{5}{2}, -\frac{5}{2})$
10. Parábola; vértice en $(\frac{5}{16}, -\frac{15}{16})$
11. Elipse; centro en $(0, 0)$
12. Elipse; centro en $(6, -4)$
13. Parábola; vértice en $(\frac{2}{25}, \frac{11}{25})$
14. Elipse; centro en $(0, 0)$
15. Parábola; vértice en $(\frac{3}{4}, \frac{3}{4})$
16. Elipse; centro en $(-1, \frac{1}{2})$
17. Hipérbola; centro en $(0, 0)$
18. Hipérbola; centro en $(-1, 2)$
19. -14

5.20 Ejercicios (pág. 174)

8. $a = \pm\frac{1}{3}\sqrt{3}$

13. (a), (b), y (e)

Capítulo 6

6.3 Ejercicios (pág. 178)

1. $y = e^{3x} - e^{2x}$
2. $y = \frac{2}{3}x^2 + \frac{1}{3}x^5$
3. $y = 4\cos x - 2\cos^2 x$
4. Cuatro veces la cantidad inicial
5. $f(x) = Cx^n$, or $f(x) = Cx^{1/n}$
6. b) $y = e^{4x} - e^{-x^3/3}$
7. $y = c_1e^{2x} + c_2e^{-2x}$
8. $y = c_1 \cos 2x + c_2 \operatorname{sen} 2x$
9. $y = e^{x}(c_1 \cos 2x + c_2 \operatorname{sen} 2x)$
10. $y = e^{-x}(c_1 + c_2x)$
11. $k = n^2\pi^2$; $f_k(x) = C \operatorname{sen} n\pi x$ $(n = 1, 2, 3, \ldots)$
13. a) $y'' - y = 0$
    b) $y'' - 4y' + 4y = 0$
    c) $y'' + y' + \frac{5}{4}y = 0$
    d) $y'' + 4y = 0$
    e) $y'' - y = 0$

14. $y = \frac{1}{3}\sqrt{6}$, $y'' = -12y = -4\sqrt{6}$

6.9 Ejercicios (pág. 190)

1. $y = c_1 + c_2e^{-x} + c_3e^{3x}$
2. $y = c_1 + c_2e^{x} + c_3e^{-x}$
3. $y = c_1 + (c_2 + c_3x)e^{-2x}$
4. $y = (c_1 + c_2x + c_3x^2)e^{x}$
5. $y = (c_1 + c_2x + c_3x^2 + c_4x^3)e^{-x}$
6. $y = c_1e^{2x} + c_2e^{-2x} + c_3 \cos 2x + c_4 \operatorname{sen} 2x$
7. $y = e^{\sqrt{2}x}(c_1 \cos \sqrt{2}x + c_2 \operatorname{sen} \sqrt{2}x) + e^{-\sqrt{2}x}(c_3 \cos \sqrt{2}x + c_4 \operatorname{sen} \sqrt{2}x)$
8. $y = c_1e^{x} + e^{-x/2}(c_2 \cos \frac{1}{2}\sqrt{3}x + c_3 \operatorname{sen} \frac{1}{2}\sqrt{3}x)$
9. $y = e^{-x}[(c_1 + c_2x) \cos x + (c_3 + c_4x) \operatorname{sen} x]$
10. $y = (c_1 + c_2x) \cos x + (c_3 + c_4x) \operatorname{sen} x$
11. $y = c_1 + c_2x + (c_3 + c_4x) \cos \sqrt{2}x + (c_5 + c_6x) \operatorname{sen} \sqrt{2}x$
12. $y = c_1 + c_2x + (c_3 + c_4x) \cos 2x + (c_5 + c_6x) \operatorname{sen} 2x$

13. $f(x) = \dfrac{1}{2m^2}(e^{mx} - \cos mx - \operatorname{sen} mx)$

15. a) $y^{(4)} - 5y'' + 4y = 0$
 b) $y''' + 6y'' + 12y' + 8y = 0$
 c) $y^{(4)} - 2y''' + y'' = 0$
 d) $y^{(4)} - 2y''' + y'' = 0$
 e) $y^{(5)} - 2y^{(4)} + y''' = 0$
 f) $y^{(4)} + 8y''' + 33y'' + 68y' + 52y = 0$
 g) $y^{(4)} - 2y'' + y = 0$
 h) $y^{(6)} + 4y'' = 0$

6.15 Ejercicios (pág. 204)

1. $y_1 = -2x - x^2 - \frac{1}{3}x^3$
2. $y_1 = \frac{1}{4}xe^{2x}$
3. $y_1 = (x - \frac{4}{3})e^{x}$
4. $y_1 = \frac{1}{3} \operatorname{sen} x$
5. $y_1 = \frac{1}{2}x^2e^{x} + e^{2x}$
6. $y_1 = \frac{1}{2}xe^{x}$
7. $y_1 = x \cosh x$
8. $y_1 = \frac{1}{24}x^4e^{-x}$
9. $50y_1 = (11 - 5x)e^{x} \operatorname{sen} 2x + (2 - 10x)e^{x} \cos 2x$
10. $y_1 = -(\frac{5}{8}x + \frac{3}{8}x^2 + \frac{1}{12}x^3)e^{-x}$

12. $y_1 = \dfrac{x^m e^{\alpha x}}{p_A^{(m)}(\alpha)}$

15. b) $2D$ c) $3D^2$ d) nD^{n-1}

16. $y = Ae^{x} + Be^{-x} + \dfrac{1}{2}e^{x}\displaystyle\int \frac{e^{-x}}{x}\,dx - \frac{1}{2}e^{-x}\int \frac{e^{x}}{x}\,dx$

17. $y = (A + \frac{1}{2}x) \operatorname{sen} 2x + (B + \frac{1}{4} \log |\cos 2x|) \cos 2x$
18. $y = Ae^{x} + Be^{-x} + \frac{1}{2} \sec x$
19. $y = (A + Bx)e^{x} + e^{e^x} - xe^{x}\displaystyle\int e^{e^x}\,dx + e^{x}\int xe^{e^x}\,dx$

20. $y = -\frac{1}{8}\log|x| + \frac{1}{3}e^{x}\int\frac{e^{-x}}{x}\,dx - \frac{1}{4}e^{2x}\int\frac{e^{-2x}}{x}\,dx$

$$+\frac{1}{24}e^{4x}\int\frac{e^{-4x}}{x}\,dx + Ae^{x} + Be^{2x} + Ce^{4x}$$

6.16 Miscelánea de ejercicios sobre ecuaciones diferenciales lineales (pág. 206)

1. $u(x) = 6(e^{4x} - e^{-x})/5\,;\quad v(x) = e^{x} - e^{-5x}$
2. $u(x) = \frac{1}{2}e^{2x-\pi}\operatorname{sen} 5x\,;\quad v(x) = \frac{5}{6}e^{-2x-\pi}\operatorname{sen} 3x$
3. $u(x) = e^{-x^2};\quad Q(x) = 4x^2 + 2$
5. $y = (A + Bx^3)e^{x} + (x^2 - 2x + 2)e^{2x}$
6. $y = Ae^{4x}\int e^{-4x-x^3/3}\,dx + Be^{4x}$
7. $y = Ax^{1/2} + Bx^{-1/2}$
8. $y = Ae^{x} + Bx^2e^{-x} - x$
9. $y = A(x^2 - 2) + B/x$
10. $y = x^{-2}[A + B(x-1)^3 + \frac{1}{9}x^3 + \frac{2}{3}x^2 - \frac{7}{6}x + \frac{1}{2} - (x-1)^3\log|x-1|]$
11. $a = 1\,,\ -1;\quad y = [Ae^{g(x)} + Be^{-g(x)}]/x$

6.21 Ejercicios (pág. 218)

2. $f(x) = u_1(x)\ (\alpha = 1)$
3. a) $A = (a - b)/2\,,\quad B = (a + b)/2$

   b) $\frac{d}{dt}\left[(t^2 - 1)\frac{dy}{dt}\right] - \alpha(\alpha + 1)y = 0\,,$ donde $\alpha = 1$ o -2, y $x = (t + 1)/2$

4. $u_1(x) = 1 + \sum_{m=1}^{\infty}(-1)^m 2^m\frac{\alpha(\alpha-2)\cdots(\alpha-2m+2)}{(2m)!}x^{2m}$ para todo x;

   $u_2(x) = x + \sum_{m=1}^{\infty}(-1)^m 2^m\frac{(\alpha-1)(\alpha-3)\cdots(\alpha-2m+1)}{(2m+1)!}x^{2m+1}$ para todo x

5. $u_1(x) = 1 + \sum_{m=1}^{\infty}\frac{(-1)^m}{(3m+2)(3m-1)\cdots 8\cdot 5}x^{3m}$ para todo x;

   $u_2(x) = x^{-2}\left(1 + \sum_{n=1}^{\infty}\frac{(-1)^n}{3^n n!}x^{3n}\right)$ para todo $x \neq 0$

6. $y = x^2\left(\frac{1}{6} + \sum_{n=1}^{\infty}\frac{(\alpha-2)(\alpha-3)\cdots(\alpha-n-1)}{n!(n+3)!}x^n\right)$ para todo x

11. b) $f(x) = \frac{1}{5}P_0(x) + \frac{4}{7}P_2(x) + \frac{8}{35}P_4(x)$

15. b) $\frac{2n}{4n^2 - 1}$

6.24 Ejercicios (pág. 231)

5. $J_{-3/2}(x) = -\left(\dfrac{2}{\pi x}\right)^{1/2}\left(\dfrac{\cos x}{x} + \operatorname{sen} x\right)$

9. a) $y = x^{1/2}[c_1 J_{1/3}(\tfrac{2}{3}x^{3/2}) + c_2 J_{-1/3}(\tfrac{2}{3}x^{3/2})]$
b) $y = x^{1/2}[c_1 J_{1/4}(\tfrac{1}{2}x^2) + c_2 J_{-1/4}(\tfrac{1}{2}x^2)]$
c) $y = x^{1/2}[c_1 J_\alpha(2\alpha x^{1+m/2}) + c_2 J_{-\alpha}(2\alpha x^{1+m/2})]$, donde $\alpha = 1/(m+2)$, con tal que $1/(m+2)$ no sea entero; si no, sustituir la J correspondiente por K
d) $y = x^{1/2}[c_1 J_\alpha(\tfrac{1}{2}x^2) + c_2 J_{-\alpha}(\tfrac{1}{2}x^2)]$, donde $\alpha = \sqrt{2}/8$

10. $y = g_\alpha$ satisface $x^2y'' + (1-2c)xy' + (a^2b^2x^{2b} + c^2 - \alpha^2b^2)y = 0$
a) $y = x^{-5/2}[c_1 J_5(2x^{1/2}) + c_2 K_5(2x^{1/2})]$
b) $y = x^{-5/2}[c_1 J_{5/2}(x) + c_2 J_{-5/2}(x)]$
c) $y = x^{-}$[illegible]$[c_1 J_1(\tfrac{2}{5}x^{5/2}) + c_2 K_1(\tfrac{2}{5}x^{5/2})]$
d) $y = x[c_1 J_0(2x^{1/2}) + c_2 K_0(2x^{1/2})]$

11. $a = 2, \quad c = 0$

12. $y = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} x^n\,; \quad y = J_0(2x^{1/2}) \quad \text{si } x > 0$

13. $b = (p_0 - a_0)/a_0, \quad c = q_0/a_0$

14. $y = x^{1/2}$

15. $t = 1: \quad y = \displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n!}{(2n)!} (2x)^n$

$t = \dfrac{1}{2}: \quad y = x^{1/2} \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{x}{2}\right)^n = x^{1/2} e^{-x/2}$

16. $u_0(x) = \cos x\,; \quad u_1(x) = \tfrac{1}{2} - \tfrac{1}{6}\cos x - \tfrac{1}{3}\cos 2x$

Capítulo 7

7.4 Ejercicios (pág. 241)

3. b) $(P^k)' = \displaystyle\sum_{m=0}^{k-1} P^m P' P^{k-1-m}$

7.12 Ejercicios (pág. 251)

1. a) $A^{-1} = 2I - A, \quad A^n = nA - (n-1)I$

b) $e^{tA} = e^t(1-t)I + te^tA = e^t \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}$

2. a) $A^{-1} = \frac{3}{2}I - \frac{1}{2}A, \quad A^n = (2^n - 1)A - (2^n - 2)I$

b) $e^{tA} = (2e^t - e^{2t})I + (e^{2t} - e^t)A = \begin{bmatrix} e^t & 0 \\ e^{2t} - e^t & e^{2t} \end{bmatrix}$

3. a) $A^{-1} = A, \quad A^n = \dfrac{1 + (-1)^n}{2} I + \dfrac{1 - (-1)^n}{2} A$

b) $e^{tA} = (\cosh t)I + (\operatorname{senh} t)A = \begin{bmatrix} \cosh t & \operatorname{senh} t \\ \operatorname{senh} t & \cosh t \end{bmatrix}$

4. a) $A^{-1} = A, \quad A^n = \dfrac{1 + (-1)^n}{2} I + \dfrac{1 - (-1)^n}{2} A$

b) $e^{tA} = (\cosh t)I + (\operatorname{senh} t)A = \begin{bmatrix} e^{-t} & 0 \\ 0 & e^t \end{bmatrix}$

5. b) $e^{tA} = e^{at}\begin{bmatrix} \cos bt & \operatorname{sen} bt \\ -\operatorname{sen} bt & \cos bt \end{bmatrix}$

7. $e^{A(t)} = I + (e - 1)A(t); \quad (e^{A(t)})' = (e - 1)A'(t) = \begin{bmatrix} 0 & e - 1 \\ 0 & 0 \end{bmatrix};$

$e^{A(t)}A'(t) = \begin{bmatrix} 0 & e \\ 0 & 0 \end{bmatrix}; \quad A'(t)e^{A(t)} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

8. a) $A^n = O \quad \text{si} \quad n \geq 3$

b) $e^{tA} = I + tA + \dfrac{1}{2}t^2A^2 = \begin{bmatrix} 1 & t & t + \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}$

9. a) $A^n = A \quad \text{si} \quad n \geq 1$

b) $e^{tA} = I + (e^t - 1)A = \begin{bmatrix} 1 & e^t - 1 & e^t - 1 \\ 0 & e^t & e^t - 1 \\ 0 & 0 & 1 \end{bmatrix}$

10. a) $A^3 = 4A^2 - 5A + 2I; \quad A^n = \begin{bmatrix} 2^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & n & 1 \end{bmatrix}$

b) $e^{tA} = \begin{bmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & te^t & e^t \end{bmatrix}$

11. $e^{tA} = I + tA + \frac{1}{2}t^2A^2$

13. $e^A e^B = \begin{bmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{bmatrix}$; $e^B e^A = \begin{bmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{bmatrix}$; $e^{A+B} = \begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$

7.15 Ejercicios (pág. 260)

1. $e^{tA} = \frac{1}{2}(3e^t - e^{3t})I + \frac{1}{2}(e^{3t} - e^t)A$
2. $e^{tA} = (\cosh \sqrt{5}\, t)I + \dfrac{1}{\sqrt{5}}(\operatorname{senh}\sqrt{5}\, t)A$
3. $e^{tA} = \frac{1}{2}e^t\{(t^2 - 2t + 2)I + (-2t^2 + 2t)A + t^2A^2\}$
4. $e^{tA} = (3e^{-t} - 3e^{-2t} + e^{-3t})I + (\frac{5}{2}e^{-t} - 4e^{-2t} + \frac{3}{2}e^{-3t})A + (\frac{1}{2}e^{-t} - e^{-2t} + \frac{1}{2}e^{-3t})A^2$
5. $e^{tA} = (4e^t - 3e^{2t} + 2te^{2t})I + (4e^{2t} - 3te^{2t} - 4e^t)A + (e^t - e^{2t} + te^{2t})A^2$
6. $e^{tA} = (4e^t - 6e^{2t} + 4e^{3t} - e^{4t})I + (-\frac{13}{3}e^t + \frac{19}{2}e^{2t} - 7e^{3t} + \frac{11}{6}e^{4t})A$
   $+ (\frac{3}{2}e^t - 4e^{2t} + \frac{7}{2}e^{3t} - e^{4t})A^2 + (-\frac{1}{6}e^t + \frac{1}{2}e^{2t} - \frac{1}{2}e^{3t} + \frac{1}{6}e^{4t})A^3$
7. b) $e^{tA} = \frac{1}{6}e^{\lambda t}\{(6 - 6\lambda t + 3\lambda^2t^2 - \lambda^3t^3)I + (6t - 6\lambda t^2 + 3\lambda^2t^3)A + (3t^2 - 3\lambda t^3)A^2$
   $+ t^3A^3\}$
8. $y_1 = c_1 \cosh \sqrt{5}\, t + \dfrac{c_1 + 2c_2}{\sqrt{5}} \operatorname{senh} \sqrt{5}\, t$, $y_2 = c_2 \cosh \sqrt{5}\, t + \dfrac{2c_1 - c_2}{\sqrt{5}} \operatorname{senh} \sqrt{5}\, t$
9. $y_1 = e^t(\cos 3t - \operatorname{sen} 3t)$, $y_2 = e^t(\cos 3t - 3 \operatorname{sen} 3t)$
10. $y_1 = e^{2t} + 4te^{2t}$, $y_2 = -2e^t + e^{2t} + 4te^{2t}$, $y_3 = -2e^t + 4e^{2t}$
11. $y_1 = c_1e^{2t}$, $y_2 = c_2e^t$, $y_3 = (c_2t + c_3)e^t$
12. $y_1 = 3e^{-t} - 3e^{-2t} + e^{-3t}$, $y_2 = -3e^{-t} + 6e^{-2t} - 3e^{-3t}$, $y_3 = 3e^{-t} - 12e^{-2t} + 9e^{-3t}$
13. $y_1 = e^{5t} + 7e^{-3t}$, $y_2 = 2e^{5t} - 2e^{-3t}$, $y_3 = -e^{5t} + e^{-3t}$
14. $y_1 = -\frac{1}{2}e^t + e^{2t} + \frac{1}{2}e^{3t}$, $y_2 = e^{2t} + e^{3t}$, $y_3 = e^{3t}$, $y_4 = e^{4t}$
15. $y_1 = 2e^{2t} - 1$, $y_2 = 2e^{2t} - t - 2$, $y_3 = 2e^{2t}$, $y_4 = e^{2t}$

7.17 Ejercicios (pág. 264)

2. c) $y_1 = (b - 1)e^x + 2(c + 1 - b)xe^x + 1$, $y_2 = ce^x + 2(c + 1 - b)xe^x$
4. $y_1 = -\frac{1}{3}e^t - \frac{1}{6}e^{4t} + \frac{1}{2}e^{2t}$, $y_2 = \frac{2}{3}e^t - \frac{1}{6}e^{4t} + \frac{1}{2}e^{2t}$
5. a) $B_0 = B$, $B_1 = AB$, $B_2 = \dfrac{1}{2!}A^2B, \ldots,$ $B_m = \dfrac{1}{m!}A^mB$
   b) $B = -m!\,(A^{-1})^{m+1}C$
6. a) $Y(t) = \left(I + tA + \dfrac{1}{2}t^2A^2 + \dfrac{1}{6}t^3A^3\right)B$, donde $B = -6A^{-4}C = -\dfrac{3}{128}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

   Esto nos da la solución particular $y_1 = y_2 = -\frac{3}{128} - \frac{3}{32}t - \frac{3}{16}t^2 - \frac{1}{4}t^3$
   b) $y_1 = y_2 = -\frac{3}{128} - \frac{3}{32}t - \frac{3}{16}t^2 - \frac{1}{4}t^3 + \frac{131}{128}e^{4t}$
7. $E = B$, $F = \dfrac{1}{\alpha}(AB + C)$
8. a) $y_1 = -\cos 2t - \frac{1}{2}\operatorname{sen} 2t$, $y_2 = -\frac{1}{2}\operatorname{sen} 2t$

b) $y_1 = 2\cosh 2t + \frac{5}{2}\operatorname{senh} 2t - \cos 2t - \frac{1}{2}\operatorname{sen} 2t$, $y_2 = \cosh 2t + \frac{1}{2}\operatorname{senh} 2t - \frac{1}{2}\operatorname{sen} 2t$

9. $y_1(x) = e^{2x} + e^{3x} - e^{x}$, $y_2(x) = -2e^{2x} - e^{3x} + 3e^{x}$

10. $y_1(x) = \frac{4}{25}e^{x} - \frac{1}{36}e^{2x} + (c_1 - \frac{119}{900}e^{-4x} + (\frac{11}{30} - c_1 - c_2)xe^{-4x}$,
$y_2(x) = \frac{1}{25}e^{x} + \frac{7}{36}e^{2x} + (c_2 - \frac{211}{900})e^{-4x} + (c_1 + c_2 - \frac{11}{30})xe^{-4x}$

11. $y_1(x) = e^{-4x}(2\cos x + \operatorname{sen} x) + \frac{31}{26}e^{x} - \frac{93}{17}$, $y_2(x) = e^{-4x}(\operatorname{sen} x + 3\cos x) - \frac{2}{13}e^{x} + \frac{6}{17}$

12. $y_1(x) = e^{-x}(x^2 + 2x + 3) + x^2 - 3x + 3$, $y_2(x) = e^{-x}(-2x - 2) + x$,
$y_3(x) = 2e^{-x} + x - 1$

7.20 Ejercicios (pág. 272)

4. c) $Y(x) = e^{x}e^{\frac{1}{2}x^2 A}B$

5. Si $A(x) = \sum_{k=0}^{\infty} x^k A_k$, entonces $Y(x) = B + xC + \sum_{k=2}^{\infty} x^k B_k$,

donde $(k+2)(k+1)B_{k+2} = \sum_{r=0}^{k} A_r B_{k-r}$ para $k \geq 0$.

7.24 Ejercicios (pág. 283)

1. a) $Y(x) = e^{x}$

b) $Y_n(x) = 2e^{x} - \sum_{k=0}^{n} \frac{x^k}{k!}$ si n es impar ; $Y_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$ si n es par

2. $Y_3(x) = \frac{x^2}{2} + \frac{x^5}{20} + \frac{x^8}{160} + \frac{x^{11}}{4400}$

3. $Y_3(x) = x + \frac{x^4}{4} + \frac{x^7}{14} + \frac{x^{10}}{160}$

4. $Y_3(x) = \frac{x^3}{3} + \frac{4x^7}{63} + \frac{8x^9}{405} + \frac{184x^{11}}{51975} + \frac{4x^{13}}{12285}$

5. a) $Y_2(x) = 1 + x + x^2 + \frac{2x^3}{3} + \frac{x^4}{6} + \frac{2x^5}{15} + \frac{x^7}{63}$

b) $M = 2$; $c = \frac{1}{2}$

c) $Y(x) = 1 + x + x^2 + \frac{4x^3}{3} + \frac{7x^4}{6} + \frac{6x^5}{5} + \cdots$

6. a) $Y_4(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{38x^9}{2835} + \frac{134x^{11}}{51975} + \frac{4x^{13}}{12285} + \frac{x^{15}}{59535}$

d) $Y(x) = \operatorname{tg} x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{62x^9}{2835} + \cdots$ para $|x| < \frac{\pi}{2}$

8. $Y_3(x) = 2 + x^2 + x^3 + \frac{3x^5}{20} + \frac{x^6}{10}$; $Z_3(x) = 3x^2 + \frac{3x^4}{4} + \frac{6x^5}{5} + \frac{3x^7}{28} + \frac{3x^8}{40}$

9. $Y_3(x) = 5 + x + \frac{x^4}{12} + \frac{x^6}{6} + \frac{2x^7}{63} + \frac{x^9}{72}$;

$$Z_3(x) = 1 + \frac{x^3}{3} + x^5 + \frac{2x^6}{9} + \frac{x^8}{8} + \frac{11x^9}{324} + \frac{7x^{11}}{264}$$

10. d) $Y_n(x) = 0$; $\lim_{n\to\infty} Y_n(x) = 0$

e) $Y_n(x) = \begin{cases} x^2 & \text{si } x \geq 0 \\ -x^2 & \text{si } x \leq 0 \end{cases}$; $\lim_{n\to\infty} Y_n(x) = \begin{cases} x^2 & \text{si } x \geq 0 \\ -x^2 & \text{si } x \leq 0 \end{cases}$

f) $Y_n(x) = \frac{2x^n}{3^n}$; $\lim_{n\to\infty} Y_n(x) = 0$

g) $Y_n(x) = \begin{cases} x^2 & \text{si } x \geq 0 \\ -x^2 & \text{si } x \leq 0 \end{cases}$; $\lim_{n\to\infty} Y_n(x) = \begin{cases} x^2 & \text{si } x \geq 0 \\ -x^2 & \text{si } x \leq 0 \end{cases}$

Capítulo 8

8.3 Ejercicios (pág. 300)

2. Todos abiertos excepto d), e), h), y j)
3. Todos abiertos excepto d)
5. e) Un ejemplo es la colección de todas las 2-esferas $B(O; 1/k)$, donde $k=1, 2, 3, \ldots$
6. a) Abierto y cerrado b) Abierto y cerrado c) Cerrado d) Abierto e) Cerrado f) Ni abierto ni cerrado g) Cerrado h) Ni abierto ni cerrado i) cerrado j) cerrado k) ni abierto ni cerrado l) Cerrado
8. e) Un ejemplo es la reunión de todos los conjuntos de la forma $S_k = \{x \mid \|x\| \leq 1 - 1/k\}$ para $k=1, 2, 3, \ldots$ Su reunión es la esfera abierta $B(O; 1)$
10. No

8.5 Ejercicios (pág. 306)

1. a) Todo (x, y)
   b) Todo $(x, y) \neq (0, 0)$
   c) Todo (x, y) con $y \neq 0$
   d) Todo (x, y) con $y \neq 0$ y $\frac{x^2}{y} \neq \frac{\pi}{2} + k\pi$ $(k = 0, 1, 2, \ldots)$
   e) Todo (x, y) con $x \neq 0$
   f) Todo $(x, y) \neq (0, 0)$
   g) Todo (x, y) con $xy \neq 1$
   h) Todo $(x, y) \neq (0, 0)$
   i) Todo $(x, y) \neq (0, 0)$
   j) Todo (x, y) con $y \neq 0$ y $0 \leq x \leq y$ o $y \leq x \leq 0$
5. $\lim_{y\to 0} f(x, y)$ no existe si $x \neq 0$
6. $(1 - m^2)/(1 + m^2)$; No
7. $y = \frac{1}{2}x^2$; f no es continua en $(0, 0)$

8. $f(0, 0) = 1$

8.9 Ejercicios (pág. 312)

1. $f'(\boldsymbol{x}; \boldsymbol{y}) = \boldsymbol{a} \cdot \boldsymbol{y}$
2. a) $f'(\boldsymbol{x}; \boldsymbol{y}) = 4\,\|\boldsymbol{x}\|^2\, \boldsymbol{x} \cdot \boldsymbol{y}$
   b) Todos los puntos de la recta $2x + 3y = \frac{3}{26}$
   c) Todos los puntos del plano $x + 2y + 3z = 0$
3. $f'(\boldsymbol{x}; \boldsymbol{y}) = \boldsymbol{x} \cdot T(\boldsymbol{y}) + \boldsymbol{y} \cdot T(\boldsymbol{x})$
4. $\dfrac{\partial f}{\partial x} = 2x + y^3 \cos(xy)$; $\dfrac{\partial f}{\partial y} = 2y \operatorname{sen}(xy) + xy^2 \cos(xy)$
5. $\dfrac{\partial f}{\partial x} = x/(x^2 + y^2)^{1/2}$; $\dfrac{\partial f}{\partial y} = y/(x^2 + y^2)^{1/2}$
6. $\dfrac{\partial f}{\partial x} = y^2/(x^2 + y^2)^{3/2}$; $\dfrac{\partial f}{\partial y} = -xy/(x^2 + y^2)^{3/2}$
7. $\dfrac{\partial f}{\partial x} = -2y/(x - y)^2$; $\dfrac{\partial f}{\partial y} = 2x/(x - y)^2$
8. $D_k f(\boldsymbol{x}) = a_k$, donde $\boldsymbol{a} = (a_1, \ldots, a_n)$
9. $D_k f(\boldsymbol{x}) = 2 \sum_{j=1}^{n} a_{kj} x_j$
10. $\dfrac{\partial f}{\partial x} = 4x^3 - 8xy^2$; $\dfrac{\partial f}{\partial y} = 4y^3 - 8x^2 y$
11. $\dfrac{\partial f}{\partial x} = \dfrac{2x}{x^2 + y^2}$; $\dfrac{\partial f}{\partial y} = \dfrac{2y}{x^2 + y^2}$
12. $\dfrac{\partial f}{\partial x} = -\dfrac{2x}{y} \operatorname{sen}(x^2)$; $\dfrac{\partial f}{\partial y} = -\dfrac{1}{y^2} \cos(x^2)$
13. $\dfrac{\partial f}{\partial x} = \dfrac{2x}{y} \sec^2 \dfrac{x^2}{y}$; $\dfrac{\partial f}{\partial y} = -\dfrac{x^2}{y^2} \sec^2 \dfrac{x^2}{y}$
14. $\dfrac{\partial f}{\partial x} = -\dfrac{y}{x^2 + y^2}$; $\dfrac{\partial f}{\partial y} = \dfrac{1}{x^2 + y^2}$
15. $\dfrac{\partial f}{\partial x} = \dfrac{1 + y^2}{1 + x^2 + y^2 + x^2 y^2}$; $\dfrac{\partial f}{\partial y} = \dfrac{1 + x^2}{1 + x^2 + y^2 + x^2 y^2}$
16. $\dfrac{\partial f}{\partial x} = y^2 x^{y^2 - 1}$; $\dfrac{\partial f}{\partial y} = 2y x^{y^2} \log x$
17. $\dfrac{\partial f}{\partial x} = -\dfrac{1}{2\sqrt{x(y - x)}}$; $\dfrac{\partial f}{\partial y} = \dfrac{\sqrt{x}}{2y\sqrt{y - x}}$
18. $n = -\frac{3}{2}$

19. $a = b = 1$
22. b) Un ejemplo es $f(x) = \boldsymbol{x} \cdot \boldsymbol{y}$, donde $\boldsymbol{y}$ es un vector dado no nulo

8.14 Ejercicios (pág. 320)

1. a) $(2x + y^3 \cos xy)\boldsymbol{i} + (2y \operatorname{sen} xy + xy^2 \cos xy)\boldsymbol{j}$
   b) $e^x \cos y\boldsymbol{i} - e^x \operatorname{sen} y\boldsymbol{j}$
   c) $2xy^3z^4\boldsymbol{i} + 3x^2y^2z^4\boldsymbol{j} + 4x^2y^3z^3\boldsymbol{k}$
   d) $2x\boldsymbol{i} - 2y\boldsymbol{j} + 4z\boldsymbol{k}$
   e) $\dfrac{2x}{x^2 + 2y^2 - 3z^2}\boldsymbol{i} + \dfrac{4y}{x^2 + 2y^2 - 3z^2}\boldsymbol{j} - \dfrac{6z}{x^2 + 2y^2 - 3z^2}\boldsymbol{k}$
   f) $y^z x^{y^z-1}\boldsymbol{i} + zy^{z-1}x^{y^z} \log x\boldsymbol{j} + y^z x^{y^z} \log x \log y\boldsymbol{k}$
2. a) $-2/\sqrt{6}$
   b) $1/\sqrt{6}$
3. $(1, 0)$, en la dirección de $\boldsymbol{i}$; $(-1, 0)$, en la dirección de $-\boldsymbol{i}$
4. $2\boldsymbol{i} + 2\boldsymbol{j}$; $\frac{14}{5}$
5. $(a, b, c) = (6, 24, -8)$ o $(-6, -24, 8)$
6. El conjunto de puntos (x, y) de la recta $5x - 3y = 6$; $\nabla f(\boldsymbol{a}) = 5\boldsymbol{i} - 3\boldsymbol{j}$
8. c) Sí
   d) $f(x, y, z) = \frac{1}{2}(x^2 + y^2 + z^2)$
11. b) implica a) y c); d) implica a), b), y c); f) implica a)

8.17 Ejercicios (pág. 327)

1. b) $F''(t) = \dfrac{\partial^2 f}{\partial x^2}[X'(t)]^2 + 2\dfrac{\partial^2 f}{\partial x\,\partial y}X'(t)Y'(t) + \dfrac{\partial^2 f}{\partial y^2}[Y'(t)]^2 + \dfrac{\partial f}{\partial x}X''(t) + \dfrac{\partial f}{\partial y}Y''(t)$
2. a) $F'(t) = 4t^3 + 2t$; $F''(t) = 12t^2 + 2$
   b) $F'(t) = (2\cos^2 t - 1)e^{\cos t \operatorname{sen} t}\cos(\cos t \operatorname{sen}^2 t) + (3\operatorname{sen}^3 t - 2\operatorname{sen} t)e^{\cos t \operatorname{sen} t}\operatorname{sen}(\cos t \operatorname{sen}^2 t)$;
   $F''(t) = (5\cos^6 t - 3\cos^4 t - 4\cos^3 t - \cos^2 t - 4\cos t)e^{\cos t \operatorname{sen} t}\cos(\cos t \operatorname{sen}^2 t)$
   $+ (14\operatorname{sen}^3 t - 12\operatorname{sen}^5 t - 4\operatorname{sen} t + 7\cos t - 9\cos^3 t)e^{\cos t \operatorname{sen} t}\operatorname{sen}(\cos t \operatorname{sen}^2 t)$
   c) $F'(t) = \dfrac{2e^{2t}\exp(e^{2t})}{1 + \exp(e^{2t})} + \dfrac{2e^{-2t}\exp(e^{-2t})}{1 + \exp(e^{-2t})}$, donde $\exp(u) = e^u$;
   $$F''(t) = \frac{4[1 + e^{2t} + \exp(e^{2t})]e^{2t}\exp(e^{2t})}{[1 + \exp(e^{2t})]^2} - \frac{4[1 + e^{-2t} + \exp(e^{-2t})]e^{-2t}\exp(e^{-2t})}{[1 + \exp(e^{-2t})]^2}$$
3. a) $-\frac{2}{3}$
   b) $x^2 - y^2$
   c) 0
4. a) $(1 + 3x^2 + 3y^2)(x\boldsymbol{i} + y\boldsymbol{j}) - (x^2 + y^2)^{1/2}\boldsymbol{k}$, o cualquier producto de éste por un escalar
   b) $\cos\theta = -[1 + (1 + 3(x^2 + y^2))^2]^{-1/2}$; $\cos\theta \to -\frac{1}{2}\sqrt{2}$ cuando $(x, y, z) \to (0, 0, 0)$

5. $U(x, y) = \frac{1}{2}\log(x^2 + y^2)$; $V(x, y) = \arctan(y/x)$
6. b) No
8. $x/x_0 + y/y_0 + z/z_0 = 3$
9. $x + y + 2z = 4$, $x - y - z = -1$
10. $c = \pm\sqrt{3}$

8.22 Ejercicios (pág. 336)

1. b) $\dfrac{\partial f}{\partial x} = -2x\,\text{sen}\,(x^2 + y^2)\cos[\cos(x^2 + y^2)]e^{\text{sen}\,[\cos(x^2+y^2)]}$

2. $\dfrac{\partial F}{\partial x} = \dfrac{1}{2}\dfrac{\partial f}{\partial u} + \dfrac{1}{2}\dfrac{\partial f}{\partial v}$; $\dfrac{\partial F}{\partial y} = -\dfrac{1}{2}\dfrac{\partial f}{\partial u} + \dfrac{1}{2}\dfrac{\partial f}{\partial v}$

3. a) $\dfrac{\partial F}{\partial s} = \dfrac{\partial f}{\partial x}\dfrac{\partial X}{\partial s} + \dfrac{\partial f}{\partial y}\dfrac{\partial Y}{\partial s}$; $\dfrac{\partial F}{\partial t} = \dfrac{\partial f}{\partial x}\dfrac{\partial X}{\partial t} + \dfrac{\partial f}{\partial y}\dfrac{\partial Y}{\partial t}$

c) $\dfrac{\partial^2 F}{\partial s\,\partial t} = \dfrac{\partial^2 f}{\partial x^2}\dfrac{\partial X}{\partial s}\dfrac{\partial X}{\partial t} + \dfrac{\partial^2 f}{\partial x\,\partial y}\left(\dfrac{\partial X}{\partial s}\dfrac{\partial Y}{\partial t} + \dfrac{\partial X}{\partial t}\dfrac{\partial Y}{\partial s}\right) + \dfrac{\partial^2 f}{\partial y^2}\dfrac{\partial Y}{\partial s}\dfrac{\partial Y}{\partial t} + \dfrac{\partial f}{\partial x}\dfrac{\partial^2 X}{\partial s\,\partial t} + \dfrac{\partial f}{\partial y}\dfrac{\partial^2 Y}{\partial s\,\partial t}$

4. a) $\dfrac{\partial F}{\partial s} = \dfrac{\partial f}{\partial x} + t\dfrac{\partial f}{\partial y}$; $\dfrac{\partial F}{\partial t} = \dfrac{\partial f}{\partial x} + s\dfrac{\partial f}{\partial y}$; $\dfrac{\partial^2 F}{\partial s^2} = \dfrac{\partial^2 f}{\partial x^2} + 2t\dfrac{\partial^2 f}{\partial x\,\partial y} + t^2\dfrac{\partial^2 f}{\partial y^2}$;

$\dfrac{\partial^2 F}{\partial t^2} = \dfrac{\partial^2 f}{\partial x^2} + 2s\dfrac{\partial^2 f}{\partial x\,\partial y} + s^2\dfrac{\partial^2 f}{\partial y^2}$; $\dfrac{\partial^2 F}{\partial s\,\partial t} = \dfrac{\partial^2 f}{\partial x^2} + (s + t)\dfrac{\partial^2 f}{\partial x\,\partial y} + st\dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial f}{\partial y}$

b) $\dfrac{\partial F}{\partial s} = t\dfrac{\partial f}{\partial x} + \dfrac{1}{t}\dfrac{\partial f}{\partial y}$; $\dfrac{\partial F}{\partial t} = s\dfrac{\partial f}{\partial x} - \dfrac{s}{t^2}\dfrac{\partial f}{\partial y}$; $\dfrac{\partial^2 F}{\partial s^2} = t^2\dfrac{\partial^2 f}{\partial x^2} + 2\dfrac{\partial^2 f}{\partial x\,\partial y} + \dfrac{1}{t^2}\dfrac{\partial^2 f}{\partial y^2}$;

$\dfrac{\partial^2 F}{\partial t^2} = s^2\dfrac{\partial^2 f}{\partial x^2} - 2\dfrac{s^2}{t^2}\dfrac{\partial^2 f}{\partial x\,\partial y} + \dfrac{s^2}{t^4}\dfrac{\partial^2 f}{\partial y^2} + \dfrac{2s}{t^3}\dfrac{\partial f}{\partial y}$; $\dfrac{\partial^2 F}{\partial s\,\partial t} = st\dfrac{\partial^2 f}{\partial x^2} - \dfrac{s}{t^3}\dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial f}{\partial x} - \dfrac{1}{t^2}\dfrac{\partial f}{\partial y}$

c) $\dfrac{\partial F}{\partial s} = \dfrac{1}{2}\dfrac{\partial f}{\partial x} + \dfrac{1}{2}\dfrac{\partial f}{\partial y}$; $\dfrac{\partial F}{\partial t} = -\dfrac{1}{2}\dfrac{\partial f}{\partial x} + \dfrac{1}{2}\dfrac{\partial f}{\partial y}$; $\dfrac{\partial^2 F}{\partial s\,\partial t} = -\dfrac{1}{4}\dfrac{\partial^2 f}{\partial x^2} + \dfrac{1}{4}\dfrac{\partial^2 f}{\partial y^2}$;

$\dfrac{\partial^2 F}{\partial s^2} = \dfrac{1}{4}\dfrac{\partial^2 f}{\partial x^2} + \dfrac{1}{2}\dfrac{\partial^2 f}{\partial x\,\partial y} + \dfrac{1}{4}\dfrac{\partial^2 f}{\partial y^2}$; $\dfrac{\partial^2 F}{\partial t^2} = \dfrac{1}{4}\dfrac{\partial^2 f}{\partial x^2} - \dfrac{1}{2}\dfrac{\partial^2 f}{\partial x\,\partial y} + \dfrac{1}{4}\dfrac{\partial^2 f}{\partial y^2}$

5. $\dfrac{\partial^2 \varphi}{\partial r^2} = \cos^2\theta\,\dfrac{\partial^2 f}{\partial x^2} + \cos\theta\,\text{sen}\,\theta\left(\dfrac{\partial^2 f}{\partial x\,\partial y} + \dfrac{\partial^2 f}{\partial y\,\partial x}\right) + \text{sen}^2\,\theta\,\dfrac{\partial^2 f}{\partial y^2}$;

$$\frac{\partial^2 \varphi}{\partial r\,\partial \theta} = -r\cos\theta\,\text{sen}\,\theta\,\frac{\partial^2 f}{\partial x^2} + r\cos^2\theta\,\frac{\partial^2 f}{\partial x\,\partial y} - r\,\text{sen}^2\,\theta\,\frac{\partial^2 f}{\partial y\,\partial x} + r\cos\theta\,\text{sen}\,\theta\,\frac{\partial^2 f}{\partial y^2}$$

$$-\text{sen}\,\theta\,\frac{\partial f}{\partial x} + \cos\theta\,\frac{\partial f}{\partial y};$$

6. $\frac{\partial F}{\partial r}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial r}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial r}+\frac{\partial f}{\partial z}\frac{\partial Z}{\partial r}$; $\frac{\partial F}{\partial s}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial s}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial s}+\frac{\partial f}{\partial z}\frac{\partial Z}{\partial s}$;

$\frac{\partial F}{\partial t}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial t}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial t}+\frac{\partial f}{\partial z}\frac{\partial Z}{\partial t}$

7. a) $\frac{\partial F}{\partial r}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}+2\frac{\partial f}{\partial z}$; $\frac{\partial F}{\partial s}=\frac{\partial f}{\partial x}-2\frac{\partial f}{\partial y}+\frac{\partial f}{\partial z}$; $\frac{\partial F}{\partial t}=\frac{\partial f}{\partial x}+3\frac{\partial f}{\partial y}-\frac{\partial f}{\partial z}$

b) $\frac{\partial F}{\partial r}=2r\left(\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}+\frac{\partial f}{\partial z}\right)$; $\frac{\partial F}{\partial s}=2s\left(\frac{\partial f}{\partial x}-\frac{\partial f}{\partial y}-\frac{\partial f}{\partial z}\right)$; $\frac{\partial F}{\partial t}=2t\left(\frac{\partial f}{\partial x}-\frac{\partial f}{\partial y}+\frac{\partial f}{\partial z}\right)$

8. $\frac{\partial F}{\partial s}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial s}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial s}+\frac{\partial f}{\partial z}\frac{\partial Z}{\partial s}$; $\frac{\partial F}{\partial t}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial t}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial t}+\frac{\partial f}{\partial z}\frac{\partial Z}{\partial t}$

9. a) $\frac{\partial F}{\partial s}=2s\frac{\partial f}{\partial x}+2s\frac{\partial f}{\partial y}+2t\frac{\partial f}{\partial z}$; $\frac{\partial F}{\partial t}=2t\frac{\partial f}{\partial x}-2t\frac{\partial f}{\partial y}+2s\frac{\partial f}{\partial z}$

b) $\frac{\partial F}{\partial s}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}+t\frac{\partial f}{\partial z}$; $\frac{\partial F}{\partial t}=\frac{\partial f}{\partial x}-\frac{\partial f}{\partial y}+s\frac{\partial f}{\partial z}$

10. $\frac{\partial F}{\partial r}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial r}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial r}$; $\frac{\partial F}{\partial s}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial s}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial s}$; $\frac{\partial F}{\partial t}=\frac{\partial f}{\partial x}\frac{\partial X}{\partial t}+\frac{\partial f}{\partial y}\frac{\partial Y}{\partial t}$

11. a) $\frac{\partial F}{\partial r}=\frac{\partial f}{\partial x}$; $\frac{\partial F}{\partial s}=\frac{\partial f}{\partial x}$; $\frac{\partial F}{\partial t}=\frac{\partial f}{\partial y}$

b) $\frac{\partial F}{\partial r}=\frac{\partial f}{\partial x}+2r\frac{\partial f}{\partial y}$; $\frac{\partial F}{\partial s}=\frac{\partial f}{\partial x}+2s\frac{\partial f}{\partial y}$; $\frac{\partial F}{\partial t}=\frac{\partial f}{\partial x}+2t\frac{\partial f}{\partial y}$

c) $\frac{\partial F}{\partial r}=\frac{1}{s}\frac{\partial f}{\partial x}$; $\frac{\partial F}{\partial s}=\frac{-r}{s^2}\frac{\partial f}{\partial x}+\frac{1}{t}\frac{\partial f}{\partial y}$; $\frac{\partial F}{\partial t}=\frac{-s}{t^2}\frac{\partial f}{\partial y}$

13. a) $\boldsymbol{f}(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$, más un vector constante cualquiera
b) $\boldsymbol{f}(x, y, z) = P(x)\boldsymbol{i} + Q(y)\boldsymbol{j} + R(z)\boldsymbol{k}$, donde P, Q, R son tres funciones cualesquiera que satisfacen $P' = p$, $Q' = q$, $R' = r$

14. a) $D\boldsymbol{f}(x,y)=\begin{bmatrix} e^{x+2y} & 2e^{x+2y} \\ 2\cos(y+2x) & \cos(y+2x)\end{bmatrix}$; $D\boldsymbol{g}(u,v,w)=\begin{bmatrix} 1 & 4v & 9w^2 \\ -2u & 2 & 0\end{bmatrix}$

b) $\boldsymbol{h}(u, v, w) = e^{u+2v^2+3w^3+4v-2u^2}\boldsymbol{i} + \operatorname{sen}(2v - u^2 + 2u + 4v^2 + 6w^3)\boldsymbol{j}$

c) $D\boldsymbol{h}(1,-1,1)=\begin{bmatrix} -3 & 0 & 9 \\ 0 & -6\cos 9 & 18\cos 9\end{bmatrix}$

15. a) $D\boldsymbol{f}(x,y,z)=\begin{bmatrix} 2x & 1 & 1 \\ 2 & 1 & 2z\end{bmatrix}$; $D\boldsymbol{g}(u,v,w)=\begin{bmatrix} v^2w^2 & 2uvw^2 & 2uv^2w \\ 0 & w^2\cos v & 2w\operatorname{sen} v \\ 2ue^v & u^2e^v & 0\end{bmatrix}$

b) $\boldsymbol{h}(u, v, w) = (u^2v^4w^4 + w^2 \operatorname{sen} v + u^2e^v)\boldsymbol{i} + (2uv^2w^2 + w^2 \operatorname{sen} v + u^4e^{2v})\boldsymbol{j}$

c) $D\boldsymbol{h}(u, 0, w) = \begin{bmatrix} 2u & w^2 + u^2 & 0 \\ 4u^3 & w^2 + 2u^4 & 0 \end{bmatrix}$

8.24 Ejercicios varios (pág. 342)

1. Un ejemplo: $f(x, y) = 3x$ cuando $x = y$, $f(x, y) = 0$ para los demás valores de x e y
2. $D_1 f(0, 0) = 0$; $D_2 f(0, 0) = -1$; $D_{2,1} f(0, 0) = 0$; $D_{1,2} f(0, 0)$ no existe
3. a) Si $\boldsymbol{a} = (a_1, a_2)$, entonces $f'(\boldsymbol{O}; \boldsymbol{a}) = a_2^3/a_1^2$ si $a_1 \neq 0$, y $f'(\boldsymbol{O}; \boldsymbol{a}) = 0$ si $a_1 = 0$
   b) La función f no es continua en el origen
4. $\dfrac{\partial f}{\partial x} = \dfrac{1}{2} e^{-xy} x^{-1/2} y^{1/2}$; $\dfrac{\partial f}{\partial y} = \dfrac{1}{2} e^{-xy} x^{1/2} y^{-1/2}$
5. $$F'''(t) = \frac{\partial^3 f}{\partial x^3}[X'(t)]^3 + 3\frac{\partial^3 f}{\partial x^2\,\partial y}[X'(t)]^2 Y'(t) + 3\frac{\partial^3 f}{\partial x\,\partial y^2} X'(t)[Y'(t)]^2$$
   $$+ \frac{\partial^3 f}{\partial y^3}[Y'(t)]^3 + 3\frac{\partial^2 f}{\partial x^2} X'(t)X''(t) + 3\frac{\partial^2 f}{\partial x\,\partial y}[X''(t)Y'(t) + X'(t)Y''(t)]$$
   $$+ 3\frac{\partial^2 f}{\partial y^2} Y'(t)Y''(t) + \frac{\partial f}{\partial x} X'''(t) + \frac{\partial f}{\partial y} Y'''(t),$$
   suponiendo que las derivadas parciales mixtas no dependen del orden de derivación
6. 8
7. a) $\dfrac{\partial g}{\partial u} = \dfrac{\partial f}{\partial x} v + \dfrac{\partial f}{\partial y} u$; $\dfrac{\partial g}{\partial v} = \dfrac{\partial f}{\partial x} u - \dfrac{\partial f}{\partial y} v$; $\dfrac{\partial^2 g}{\partial u\,\partial v} = uv\dfrac{\partial^2 f}{\partial x^2} + (u^2 - v^2)\dfrac{\partial^2 f}{\partial x\,\partial y}$
   $- uv\dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial f}{\partial x}$ (b) $a = \frac{1}{2}$, $b = -\frac{1}{2}$
10. a) $\varphi'(t) = A'(t)\int_c^{B(t)} f[A(t), y]\,dy + B'(t)\int_a^{A(t)} f[x, B(t)]\,dx$
    b) $\varphi'(t) = 2te^{t^2}(2e^{t^2} - e^a - e^c)$
13. Una esfera con centro en el origen y radio $\sqrt{2}$
14. $f(x) = x^2$

Capítulo 9

9.3 Ejercicios (pág. 349)

1. $f(x, y) = \operatorname{sen}(x - \frac{4}{3}y)$
2. $f(x, y) = e^{x+5y/2} - 1$
3. a) $u(x, y) = x^2y^2e^{xy}$
   b) $v(x, y) = 2 + \log\left|\dfrac{x}{y}\right|$

5. $A = B = C = 1$, $D = -3$; $f(x, y) = \varphi_1(3x + y) + \varphi_2(x - y)$
6. $G(x, y) = x - y$

9.8 Ejercicios (pág. 368)

1. $\partial X/\partial v = (1 + xu)/(x - y)$; $\partial Y/\partial u = (1 - yv)/(x - y)$; $\partial Y/\partial v = (1 + yu)/(y - x)$
2. $\partial X/\partial y = -(1 + xu)/(1 + u)$; $\partial V/\partial u = (1 - yv)/(1 + yu)$; $\partial V/\partial y = (1 - x)/(1 + u)$
3. $\dfrac{\partial X}{\partial v} = \dfrac{\partial(F, G)}{\partial(y, v)} \Big/ \dfrac{\partial(F, G)}{\partial(x, y)}$; $\dfrac{\partial Y}{\partial u} = \dfrac{\partial(F, G)}{\partial(u, x)} \Big/ \dfrac{\partial(F, G)}{\partial(x, y)}$; $\dfrac{\partial Y}{\partial v} = \dfrac{\partial(F, G)}{\partial(v, x)} \Big/ \dfrac{\partial(F, G)}{\partial(x, y)}$
4. $T = \pm \dfrac{1}{\sqrt{751}} (24\boldsymbol{i} - 4\sqrt{7}\,\boldsymbol{j} + 3\sqrt{7}\,\boldsymbol{k})$
5. $2\boldsymbol{i} + \boldsymbol{j} + \sqrt{3}\,\boldsymbol{k}$, o cualquier producto no nulo de éste por un escalar
6. $\partial x/\partial u = 0$, $\partial x/\partial v = \pi/12$
8. $n = 2$
9. $\partial f/\partial x = -1/(2y + 2z + 1)$; $\partial f/\partial y = -2(y + z)/(2y + 2z + 1)$;
   $\partial^2 f/(\partial x\, \partial y) = 2/(2y + 2z + 1)^3$
10. $\partial^2 z/(\partial x\, \partial y) = [\text{sen}(x + y)\cos^2(y + z) + \text{sen}(y + z)\cos^2(x + y)]/\cos^3(y + z)$
11. $\dfrac{\partial f}{\partial x} = -\dfrac{D_1F + 2xD_2F}{D_1F + 2zD_2F}$; $\dfrac{\partial f}{\partial y} = -\dfrac{D_1F + 2yD_2F}{D_1F + 2zD_2F}$
12. $D_1F = f'[x + g(y)]$; $D_2F = f'[x + g(y)]g'(y)$; $D_{1,1}F = f''[x + g(y)]$;
    $D_{1,2}F = f''[x + g(y)]g'(y)$; $D_{2,2}F = f''[x + g(y)][g'(y)]^2 + f'[x + g(y)]g''(y)$

9.13 Ejercicios (pág. 381)

1. Mínimo absoluto en $(0, 1)$
2. Punto de ensilladura en $(0, 1)$
3. Punta de ensilladura en $(0, 0)$
4. Mínimo absoluto en cada punto de la recta $y = x + 1$
5. Punto de ensilladura en $(1, 1)$
6. Mínimo absoluto en $(1, 0)$
7. Punto de ensilladura en $(0, 0)$
8. Punto de ensilladura en $(0, 6)$ y en $(x, 0)$, para todo x; mínimo relativo en $(0, y)$, $0 < y < 6$; máximo relativo en $(2, 3)$ y en $(0, y)$ para $y < 0$ e $y > 6$
9. Punto de ensilladura en $(0, 0)$; mínimo relativo en $(1, 1)$
10. Puntos de ensilladura en $(n\pi + \pi/2, 0)$, siendo n un entero
11. Mínimo absoluto en $(0, 0)$; Punto de ensilladura en $(-\frac{1}{4}, -\frac{1}{2})$
12. Mínimo absoluto en $(-\frac{1}{26}, -\frac{3}{26})$; máximo absoluto en $(1, 3)$
13. Máximo absoluto en $(\pi/3, \pi/3)$; mínimo absoluto en $(2\pi/3, 2\pi/3)$; máximo relativo en (π, π); mínimo relativo en $(0, 0)$; punto de ensilladura en $(0, \pi)$ y $(\pi, 0)$
14. Puntos de ensilladura en $(1, 1)$
15. Máximo absoluto en cada punto de la circunferencia $x^2 + y^2 = 1$; mínimo absoluto en $(0, 0)$

17. c) Máximo relativo en $(2, 2)$; no existen mínimos relativos; puntos de ensilladura en $(0, 3), (3, 0)$, y $(3, 3)$
18. Máximo relativo $\frac{1}{8}$ en $(\frac{1}{2}, \frac{1}{2})$ y $(-\frac{1}{2}, -\frac{1}{2})$; mínimo relativo $-\frac{1}{8}$ en $(\frac{1}{2}, -\frac{1}{2})$ y $(-\frac{1}{2}, \frac{1}{2})$; puntos de ensilladura en $(0, 0)$, $(\pm 1, 0)$, y $(0, \pm 1)$; máximo absoluto 1 en $(1, -1)$ y $(-1, 1)$; mínimo absoluto -1 en $(1, 1)$ y $(-1, -1)$
19. a) $a = 1$, $b = -\frac{1}{6}$
    b) $a = 6 \log 2 - 3\pi/2$, $b = \pi - 3 \log 2$
21. Sea $x^* = \frac{1}{n} \sum_{i=1}^{n} x_i$, $y^* = \frac{1}{n} \sum_{i=1}^{n} y_i$, $u_i = x_i - x^*$. Entonces $a = \left(\sum_{i=1}^{n} y_i u_i\right) \Big/ \left(\sum_{i=1}^{n} u_i^2\right)$, y $b = y^* - ax^*$
22. Sea $x^* = \frac{1}{n} \sum_{i=1}^{n} x_i$, $y^* = \frac{1}{n} \sum_{i=1}^{n} y_i$ $z^* = \frac{1}{n} \sum_{i=1}^{n} z_i$, $u_i = x_i - x^*$, $v_i = y_i - y^*$, y pongamos

$$\Delta = \begin{vmatrix} \sum u_i^2 & \sum u_i v_i \\ \sum u_i v_i & \sum v_i^2 \end{vmatrix}$$

, considerándose las sumas para $i = 1, 2, \ldots, n$. Entonces

$$a = \frac{1}{\Delta} \begin{vmatrix} \sum u_i z_i & \sum u_i v_i \\ \sum v_i z_i & \sum v_i^2 \end{vmatrix}, \quad b = \frac{1}{\Delta} \begin{vmatrix} \sum v_i z_i & \sum u_i v_i \\ \sum u_i z_i & \sum u_i^2 \end{vmatrix}, \quad c = z^* - ax^* - by^*$$

25. Autovalores 4, 16, 16; mínimo relativo en $(1, 1, 1)$

9.15 Ejercicios (pág. 387)

1. El valor máximo es $1/4$; no tiene mínimo
2. El máximo es 2; el mínimo es 1
3. a) El máximo es $\frac{\sqrt{a^2 + b^2}}{ab}$ en $(b(a^2 + b^2)^{-1/2}, a(a^2 + b^2)^{-1/2})$; el mínimo es $-\frac{\sqrt{a^2 + b^2}}{ab}$ en $(-b(a^2 + b^2)^{-1/2}, -a(a^2 + b^2)^{-1/2})$
   b) El mínimo es $a^2b^2/(a^2 + b^2)$ en $\left(\frac{ab^2}{a^2 + b^2}, \frac{a^2b}{a^2 + b^2}\right)$; no tiene máximo
4. El máximo es $1 + \sqrt{2}/2$ en los puntos $(n\pi + \pi/8, n\pi - \pi/8)$, siendo n un entero cualquiera; el mínimo es $1 - \sqrt{2}/2$ en $(n\pi + 5\pi/8, n\pi + 3\pi/8)$, siendo n cualquier entero
5. El máximo es 3 en $(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3})$; el mínimo es -3 en $(-\frac{1}{3}, \frac{2}{3}, -\frac{2}{3})$
6. $(0, 0, 1)$ y $(0, 0, -1)$
7. 1

8. (1, 0, 0), (0, 1, 0), (−1, 0, 0), (0, −1, 0)

9. $\dfrac{a^a b^b c^c}{(a+b+c)^{a+b+c}}$ at $\left(\dfrac{a}{a+b+c}, \dfrac{b}{a+b+c}, \dfrac{c}{a+b+c}\right)$

10. $abc\sqrt{3}/2$

11. $5 \log r + 3 \log \sqrt{3}$

12. $m^2 = \dfrac{A + C - \sqrt{(A-C)^2 + 4B^2}}{2(AC - B^2)}$

13. $(4 \pm \sqrt{5})/\sqrt{2}$

14. El ángulo es $\pi/3$; anchura del fondo $c/3$, el área máxima es $c^2/(4\sqrt{3})$

Capítulo 10

10.5 Ejercicios (pág. 399)

1. $-\frac{14}{15}$
2. $-2\pi a^2$
3. $\frac{1}{35}$
4. $\frac{4}{3}$
5. 0
6. 40
7. $\frac{23}{6}$
8. $\frac{5}{2}$
9. $-\frac{369}{10}$
10. -2π
11. 0
12. (a) $-2\sqrt{2}\,\pi$
    (b) $-\pi$

10.9 Ejercicios (pág. 403)

1. $\frac{23}{6}$
2. $2a^3$
3. $a = (3c/2)^{1/2}$
4. 0
5. $8\pi(\operatorname{sen}\theta - \cos\theta)$
6. $\pi a^3/4$
7. $-\sqrt{2}$
8. $256a^3/15$
9. $2\pi^2 a^3(1 + 2\pi^2)$
10. $[(2 + t_0^2)^{3/2} - 2\sqrt{2}]/3$
12. momento de inercia= $4a^4$
13. $2\pi/3$
14. $\dfrac{600 - 36\sqrt{2} - 49 \log(9 - 4\sqrt{2})}{64[6\sqrt{2} + \log(3 + 2\sqrt{2})]}$
15. $\bar{x} = \dfrac{6ab^2}{3a^2 + 4\pi^2 b^2}$; $\bar{y} = -\dfrac{6\pi ab^2}{3a^2 + 4\pi^2 b^2}$
16. $I_x = (a^2 + b^2)^{1/2}[\pi a^4 + (4\pi^3 - \pi/2)a^2b^2 + 32\pi^5 b^4/5]$
    $I_y = (a^2 + b^2)^{1/2}[\pi a^4 + (4\pi^3 + \pi/2)a^2b^2 + 32\pi^5 b^4/5]$

10.13 Ejercicios (pág. 409)

1. Todos excepto (*f*) son conexos
6. a) No conservativo
   b) $(2e^{2\pi} - 5e^{\pi} - 5\pi - 3)/10$
7. b) 3
8. $\frac{8}{5}$
10. $4b^2 - 8\pi b + 4$; se presenta el mínimo cuando $b = \pi$

10.18 Ejercicios (pág. 420)

1. $\varphi(x, y) = \frac{1}{2}(x^2 + y^2) + C$
2. $\varphi(x, y) = x^3y + C$
3. $\varphi(x, y) = x^2e^y + xy - y^2 + C$
4. $\varphi(x, y) = x \operatorname{sen} y + y \cos x + (x^2 + y^2)/2 + C$
5. $\varphi(x, y) = x \operatorname{sen}(xy) + C$
6. $\varphi(x, y, z) = (x^2 + y^2 + z^2)/2 + C$
7. $\varphi(x, y, z) = x^2/2 - y^2/2 + xz - yz + C$
8. f no es un gradiente
9. f no es un gradiente
10. f no es un gradiente
11. $\varphi(x, y, z) = y^2 \operatorname{sen} x + xz^3 - 4y + 2z + C$
12. $\varphi(x, y, z) = x + 2x^2y - x^3z^2 + 2y - z^3 + C$
13. b) $\varphi(x, y) = \dfrac{ar^{n+1}}{n+1} + C$ si $n \neq -1$; $\varphi(x, y) = a \log r + C$ si $n = -1$
15. $\varphi(x) = \dfrac{r^{p+2}}{p+2} + C$ si $p \neq -2$; $\varphi(x) = \log r + C$ si $p = -2$
16. $\varphi(x) = g(r) + C$

10.20 Ejercicios (pág. 425)

1. $x^2/2 + 2xy + y^2/2 = C$
2. $x^2y = C$
3. $x^3/3 - xy - y/2 + (\operatorname{sen} 2y)/4 = C$
4. $\cos 2x \operatorname{sen} 3y = C$
5. $x^3y + 4x^2y^2 - 12e^y + 12ye^y = C$
6. $\int Q(x)e^{\int P(x)dx}\,dx - ye^{\int P(x)dx} = C$
8. a) $x + y = Cy^2$
   b) $y^3/x^3 - 3 \log |x| = C$
9. a) $6(xy)^{1/2} - (y/x)^{3/2} = C$; $(x^5y)^{-1/2}$ es un factor integrante
   b) $x + e^{-x} \operatorname{sen} y = C$; $e^{-x} \cos y$ es un factor integrante
10. $x^3y^4 + x^4y^5 = C$, $10x^3y^4 + x^5y^5 = C$, respectivamente; x^2y^3 es un factor integrante común

Capítulo 11

11.9 Ejercicios (pág. 442)

1. $\frac{1}{3}$
2. 1
3. $2\sqrt{3} - \frac{38}{3}$
4. $\pi^2/4$
5. 2
6. 2π
7. 6
8. $t^{-3}(e^{t^2} - e^t) + t^{-2} - t^{-1}$

10. $\frac{1}{6}$
11. $\frac{1}{5}(\frac{21}{8} - \sqrt{2})$
12. $\pi/2$
13. $(\log 2)/6$

11.15 Ejercicios (pág. 453)

1. $-3\pi/2$
2. $\frac{3}{2} + \cos 1 + \operatorname{sen} 1 - \cos 2 - 2 \operatorname{sen} 2$
3. $e - e^{-1}$
4. $\frac{7}{8} \log 2$
5. $\pi^2 - \frac{40}{9}$
6. 6
7. $\frac{80}{3}$
8. a) $\frac{8}{3}$ b) 2 c) 320π
9. $\int_0^1 \left[\int_x^1 f(x, y)\, dy\right] dx$
10. $\int_0^4 \left[\int_{x/2}^{\sqrt{x}} f(x, y)\, dy\right] dx$
11. $\int_1^2 \left[\int_1^{y^2} f(x, y)\, dx\right] dy$
12. $\int_0^1 \left[\int_{2-y}^{1+\sqrt{1-y^2}} f(x, y)\, dx\right] dy$
13. $\int_{-1}^0 \left[\int_{-\sqrt{4y+4}}^{\sqrt{4y+4}} f(x, y)\, dx\right] dy + \int_0^8 \left[\int_{-\sqrt{4y+4}}^{2-y} f(x, y)\, dx\right] dy$
14. $\int_0^1 \left[\int_{e^y}^{e} f(x, y)\, dx\right] dy$
15. $\int_{-1}^0 \left[\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} f(x, y)\, dx\right] dy + \int_0^1 \left[\int_{-\sqrt{1-y}}^{\sqrt{1-y}} f(x, y)\, dx\right] dy$
16. $\int_0^1 \left[\int_{y^{1/2}}^{y^{1/3}} f(x, y)\, dx\right] dy$
17. $\int_{-1}^0 \left[\int_{-2 \operatorname{arcsen} y}^{\pi} f(x, y)\, dx\right] dy + \int_0^1 \left[\int_{\operatorname{arcsen} y}^{\pi - \operatorname{arcsen} y} f(x, y)\, dx\right] dy$
18. $\int_{-2}^0 \left[\int_{2x+4}^{4-x^2} f(x, y)\, dy\right] dx$
19. $\int_0^1 \left[\int_x^{2-x} (x^2 + y^2)\, dy\right] dx = \frac{4}{3}$
20. $y = 0, \quad y = x \operatorname{tg} c, \quad x^2 + y^2 = a^2, \quad x^2 + y^2 = b^2$
21. a) $\int_1^8 \left[\int_{y^{1/3}}^{y} f(x, y)\, dx\right] dy$ b) $4e^8 + 2e/3$

22. $m = 2$; $n = 1$

11.18 Ejercicios (pág. 461)

1. $\bar{x} = -\frac{1}{2}, \quad \bar{y} = \frac{8}{5}$
2. $\bar{x} = 1, \quad \bar{y} = 0$
3. $\bar{x} = \frac{18}{13}, \quad \bar{y} = \frac{50}{39}$
4. $\bar{x} = \pi/2, \quad \bar{y} = \pi/8$
5. $\bar{x} = (\sqrt{2} + 1)\left(\dfrac{\pi\sqrt{2}}{4} - 1\right) = \dfrac{\pi}{2} + \dfrac{\pi\sqrt{2}}{4} - 1 - \sqrt{2}, \quad \bar{y} = \dfrac{\sqrt{2} + 1}{4}$
6. $\bar{x} = \dfrac{2a^2 \log a - a^2 + 1}{4(a \log a - a + 1)}, \quad \bar{y} = \dfrac{a(\log a)^2}{2(a \log a - a + 1)} - 1$
7. $\bar{x} = \bar{y} = \frac{1}{5}$
8. $\bar{x} = \bar{y} = 256/(315\pi)$
9. $\frac{26}{3} - \frac{15}{2} \log 3$
10. $\bar{x} = \frac{2}{3} \|\overrightarrow{AB}\|, \quad \bar{y} = \frac{2}{3} \|\overrightarrow{AD}\|$; suponiendo que los ejes x e y coincidan con los lados AB y , respectivamente
11. $I_x = \dfrac{5\pi}{12}, \quad I_y = \dfrac{2\pi^3}{3} - \pi$
12. $I_x = \frac{1}{12}b^3(a - c), \quad I_y = \frac{1}{12}b(a^3 - c^3)$
13. $I_x = I_y = (1 - 5\pi/16)r^4$
14. $I_x = I_y = \frac{9}{8}$
15. $I_x = \frac{1}{64}[(4a - 1)e^{4a} - 1], \quad I_y = \frac{1}{32}[(a^3 - 3a^2 + 6a - 6)e^{2a} + 6]$
16. $I_x = \frac{72}{105}, \quad I_y = \frac{148}{45}$
19. $\frac{1}{3}h[\sqrt{2} + \log(1 + \sqrt{2})]$
20. $h^2 + \frac{1}{2}r^2$
21. a) $(\frac{13}{6}, 1)$
    b) $(\frac{7}{3}, \frac{9}{2})$
    c) $(\frac{11}{4}, \frac{11}{4})$
    d) $(\frac{19}{8}, \frac{13}{8})$
22. $h = 2\sqrt{3}$
23. $h > r\sqrt{2}$

11.22 Ejercicios (pág. 471)

1. a) -4
   b) 4
   c) 8
   d) 4π
   e) $3\pi/2$

2. 0
3. $n = 3$
4. $-\pi$
9. $g(x, y) = \pm[P^2(x, y) + Q^2(x, y)]^{1/2}$

*** 11.25 Ejercicios (pág. 478)**

1. b) 0
2. 0, 2π, -2π
3. A lo sumo tres
4. A lo sumo siete
5. a) -3
6. 2π

11.28 Ejercicios (pág. 488)

1. $\int_0^{2\pi}\left[\int_0^a f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta$
2. $\int_{-\pi/2}^{\pi/2}\left[\int_0^{2\cos\theta} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta$
3. $\int_0^{2\pi}\left[\int_a^b f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta$
4. $\int_0^{\pi/2}\left[\int_0^{g(\theta)} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta,$ donde $g(\theta) = 1/(\cos\theta + \mathrm{sen}\,\theta)$
5. $\int_0^{\pi/4}\left[\int_0^{\tan\theta\sec\theta} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta + \int_{\pi/4}^{3\pi/4}\left[\int_0^{\csc\theta} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta$
$+ \int_{3\pi/4}^{\pi}\left[\int_0^{\tan\theta\sec\theta} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta$
6. $\frac{3}{4}\pi a^4$
7. $\frac{1}{6}a^3[\sqrt{2} + \log(1 + \sqrt{2})]$
8. $\sqrt{2} - 1$
9. $\pi a^4/8$
10. $\int_0^{\pi/4}\left[\int_0^{\sec\theta} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta + \int_{\pi/4}^{\pi/2}\left[\int_0^{\csc\theta} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta$
11. $\int_{\pi/4}^{\pi/3}\left[\int_0^{2\sec\theta} f(r)r\,dr\right]d\theta$
12. $\int_0^{\pi/2}\left[\int_{g(\theta)}^1 f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta,$ donde $g(\theta) = 1/(\cos\theta + \mathrm{sen}\,\theta)$
13. $\int_0^{\pi/4}\left[\int_{\tan\theta\sec\theta}^{\sec\theta} f(r\cos\theta, r\,\mathrm{sen}\,\theta)r\,dr\right]d\theta$
14. $\pi^4/3$
15. a) $u = 7x - y,\quad v = -5x + y$
 b) 60

17. a) $1 + 2u$
c) $\frac{14}{3}$
d) $2 + \frac{2}{\sqrt{3}}\left(\arctan\frac{1}{\sqrt{3}} - \arctan\frac{5}{\sqrt{3}}\right)$

18. a) $4(u^2 + v^2)$
c) 0

19. $\frac{\pi}{1-p}[(p^2 + r^2)^{1-p} - p^{2(1-p)}]$ si $p \neq 1$; $\pi \log(1 + r^2)$ si $p = 1$.

$I(p, r)$ tiende a un límite finito cuando $p > 1$

11.34 Ejercicios (pág. 504)

1. $\frac{1}{364}$
2. $\log\sqrt{2} - \frac{5}{16}$
3. $\frac{1}{48}$
4. $\frac{4}{5}\pi abc$
5. $\pi/6$
6. $\int_0^1 \left\{\int_0^x \left[\int_0^{1-x} f(x, y, z)\, dy\right] dz + \int_x^1 \left[\int_{z-x}^{1-x} f(x, y, z)\, dy\right] dz\right\} dx$
7. $\int_0^1 \left\{\int_{-z}^{z} \left[\int_{-\sqrt{z^2-x^2}}^{\sqrt{z^2-x^2}} f(x, y, z)\, dy\right] dx\right\} dz$
8. $\int_0^1 \left\{\int_0^{x^2} \left[\int_0^1 f(x, y, z)\, dy\right] dz + \int_{x^2}^{1+x^2} \int_{\sqrt{z-x^2}}^{1} \left[f(x, y, z)\, dy\right] dz\right\} dx$
10. $16\pi/3$
11. $\frac{1}{6}$
12. $\frac{1}{60}\pi a^2 h(3a^2 + 2h^2)$
13. $\frac{4}{3}\pi a^3$
14. $\frac{4}{3}\pi(b^3 - a^3)$
15. $\frac{4}{3}\pi R^3(a^2 + b^2 + c^2)^{-1/2}$
18. $\frac{2}{3}\pi(5\sqrt{5} - 4)$
19. $\frac{32}{9}$
20. $\frac{4}{5}\pi(b^5 - a^5)$
22. En el eje, a una distancia $\frac{2}{5}h$ de la base
23. En el eje, a una distancia $\frac{1}{5}h$ de la base
24. En el eje de simetría a una distancia $\frac{3}{8} \cdot \frac{b^4 - a^4}{b^3 - a^3}$ del plano secante de los hemisferios.
25. $\bar{x} = \bar{y} = \bar{z} = \frac{7}{12}h$ (suponiendo el vértice citado en el origen)
26. $\frac{3}{20}M(a^2 + 4h^2)$
27. $\frac{2}{5}MR^2$
28. $\frac{3}{5}Ma^2$
29. $2^{1/4}$

12.4 Ejercicios (pág. 517)

1. $(a_2b_3 - a_3b_2)(x - x_0) + (a_3b_1 - a_1b_3)(y - y_0) + (a_1b_2 - a_2b_1)(z - z_0) = 0$;

$$\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} = (a_2b_3 - a_3b_2)\boldsymbol{i} + (a_3b_1 - a_1b_3)\boldsymbol{j} + (a_1b_2 - a_2b_1)\boldsymbol{k}$$

2. $x^2/a^2 + y^2/b^2 = z$; $\dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v} = -2bu^2 \cos v\boldsymbol{i} - 2au^2 \operatorname{sen} v\boldsymbol{j} + abu\boldsymbol{k}$

3. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$; $\dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v} = abc \operatorname{sen} u \left(\dfrac{\operatorname{sen} u \cos v}{a} \boldsymbol{i} + \dfrac{\operatorname{sen} u \operatorname{sen} v}{b} \boldsymbol{j} + \dfrac{\cos u}{c} \boldsymbol{k} \right)$

4. $z = f(\sqrt{x^2 + y^2})$; $\dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v} = -uf'(u) \cos v\boldsymbol{i} - uf'(u) \operatorname{sen} v\boldsymbol{j} + u\boldsymbol{k}$

5. $\dfrac{y^2}{a^2} + \dfrac{z^2}{b^2} = 1$; $\dfrac{\partial \boldsymbol{r}}{\partial u} \times \dfrac{\partial \boldsymbol{r}}{\partial v} = b \operatorname{sen} v\boldsymbol{j} + a \cos v\boldsymbol{k}$

6. $(\sqrt{x^2 + y^2} - a)^2 + z^2 = b^2$;

$$\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} = b(a + b \cos u)(\cos u \operatorname{sen} v\boldsymbol{i} + \cos u \cos v\boldsymbol{j} + \operatorname{sen} u\boldsymbol{k})$$

7. $|abc| \cosh v \left[\left(\dfrac{\operatorname{sen}^2 u}{a^2} + \dfrac{\cos^2 u}{b^2} \right) \cosh^2 v + \dfrac{\operatorname{senh}^2 v}{c^2} \right]^{1/2}$

8. $\sqrt{128v^2 + 4}$
9. $|u - v| \sqrt{36u^2v^2 + 9(u + v)^2 + 4}$
10. $\sqrt{u^4 + u^2}$

12.6 Ejercicios (pág. 524)

2. $\pi a^2\sqrt{3}$
3. $(2\pi - 4)a^2$
4. 4
5. a) Un paraboloide de revolución
   b) $-2u^2 \cos v\boldsymbol{i} - 2u^2 \operatorname{sen} v\boldsymbol{j} + u\boldsymbol{k}$
   c) $n = 6$
6. $\sqrt{2}\, \pi a^2/4$
7. $2\pi\sqrt{6}$
8. $2\pi a^2(3\sqrt{3} - 1)/3$
9. $4\pi^2 ab$
11. a) Un círculo unitario en el plano xy; un semicírculo unitario en el plano xz, con $z \leq 0$; un semicírculo unitario en el plano $x = y$ con $z \leq 0$
    b) La semiesfera $x^2 + y^2 + z^2 = 1$, $z \leq 0$

c) La esfera $x^2 + y^2 + z^2 = 1$ excepto el Polo Norte; la línea que une el Polo Norte y (x, y, z) intercepta al plano xy en $(u, v, 0)$

12.10 Ejercicios (pág. 532)

1. $4\pi/3$
3. $\bar{x} = \bar{y} = \bar{z} = a/2$
4. $\frac{1}{2}$
7. 0
8. $\pi\sqrt{2}$
9. En el eje del cono, a una distancia $\frac{1}{4}a(1 - \cos \alpha)/[1 - \cos (\alpha/2)]$ del centro de la esfera
10. $\pi a^3 h + \frac{2}{3}\pi a h^3$
11. $3\pi a^3 h + \frac{2}{3}\pi a h^3$
12. $2\pi/3$
13. $-\pi/3$

12.13 Ejercicios (pág. 539)

1. 0
2. $-\pi$
3. -4
4. $\frac{4}{3}$

12.15 Ejercicios (pág. 545)

1. a) $\operatorname{div} \boldsymbol{F}(x, y, z) = 2x + 2y + 2z$; $\operatorname{rot} \boldsymbol{F}(x, y, z) = \boldsymbol{0}$
b) $\operatorname{div} \boldsymbol{F}(x, y, z) = 0$; $\operatorname{rot} \boldsymbol{F}(x, y, z) = 2\boldsymbol{i} + 4\boldsymbol{j} + 6\boldsymbol{k}$
c) $\operatorname{div} \boldsymbol{F}(x, y, z) = -x \operatorname{sen} y$; $\operatorname{rot} \boldsymbol{F}(x, y, z) = \boldsymbol{i} + \boldsymbol{j}$
d) $\operatorname{div} \boldsymbol{F}(x, y, z) = ye^{xy} - x \operatorname{sen}(xy) - 2xz \operatorname{sen}(xz^2)$;
$\operatorname{rot} \boldsymbol{F}(x, y, z) = z^2 \operatorname{sen}(xz^2)\boldsymbol{j} - [xe^{xy} + y \operatorname{sen}(xy)]\boldsymbol{k}$
e) $\operatorname{div} \boldsymbol{F}(x, y, z) = 2x \operatorname{sen} y + 2y \operatorname{sen}(xz) - xy \operatorname{sen} z \cos(\cos z)$;
$\operatorname{rot} \boldsymbol{F}(x, y, z) = [x \operatorname{sen}(\cos z) - xy^2 \cos(xz)]\boldsymbol{i} - y \operatorname{sen}(\cos z)\boldsymbol{j} + [y^2 z \cos(xz) - x^2 \cos y]\boldsymbol{k}$
2. $\boldsymbol{0}$
4. $n = -3$
5. No hay tal campo vectorial
10. Un tal campo es $\mu(x, y, z) = (xyz)^{-2}$
11. $\operatorname{div}(\boldsymbol{V} \times \boldsymbol{r}) = 0$; $\operatorname{rot}(\boldsymbol{V} \times \boldsymbol{r}) = (c + 1)\boldsymbol{V}$
13. $16(a + b)$

*** 12.17 Ejercicios (pág. 551)**

1. $(3x - 2z)\boldsymbol{j} - x\boldsymbol{k}$ es un tal campo
2. $(x^2/2 - xy - yz + z^2/2)\boldsymbol{j} + (x^2/2 - xz)\boldsymbol{k}$ es un tal campo
3. $(x^2y/2 + z^2/2)\boldsymbol{j} + \nabla f(x, y)$ para cualquier f independiente de z
5. $\boldsymbol{G}(x, y, z) = \dfrac{yz}{r(x^2 + y^2)}\boldsymbol{i} - \dfrac{xz}{r(x^2 + y^2)}\boldsymbol{j}$ satisface $\operatorname{rot} \boldsymbol{G} = r^{-3}\boldsymbol{r}$ para todos los puntos que no estén en el eje z

6. $f(r) = Cr^{-3}$
9. $F(x, y, z) = -\frac{1}{3}(z^3\boldsymbol{i} + x^3\boldsymbol{j} + y^3\boldsymbol{k})$, $\quad G(x, y, z) = \frac{1}{3}\nabla(x^3y + y^3z + z^3x)$
10. c) $3\pi/2$

12.21 Ejercicios (pág. 563)

1. 3
2. a) 144π
   b) -16π
   c) 128π
3. a) $3\,|V|$
   b) $9\,|V|\,\bar{z}$
   c) $|V|\,\bar{x}$
   d) $4I_z$
15. 8π

Capítulo 13

13.4 Ejercicios (pág. 575)

2. $A_1 \cup A_2 \cup A_3 = (A_1 \cap A_2' \cap A_3') \cup (A_2 \cap A_3') \cup A_3$; $\quad \bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{n-1} (A_k \cap \bigcap_{j=k+1}^{n} A_j') \cup A_n$

3.

	(i)	(ii)	(iii)	(iv)	(v)
(a)	$A' \cap B'$	$A \cap B'$	$A \cup B$	$(A \cap B') \cup (A' \cap B)$	$A' \cup B'$
(b)	500	200	500	300	800

6. $\mathscr{B}_2 = \{\varnothing, A_1, A_2, A_1 \cup A_2, A_1', A_2', A_1' \cap A_2', S\}$
7. $\mathscr{B}_3 = \{\varnothing, A_1, A_2, A_3, A_1 \cup A_2, A_2 \cup A_3, A_1 \cup A_3, A_1 \cup A_2 \cup A_3, A_1', A_2', A_3', A_1' \cap A_2', A_2' \cap A_3', A_1' \cap A_3', A_1' \cap A_2' \cap A_3', S\}$ (si $n > 3$)

13.7 Ejercicios (pág. 581)

1. $A \subseteq B'$
2. $x \in A' \cap B' \cap C'$
3. $x \in A \cap B' \cap C'$
4. $x \in A \cup B \cup C$
5. $x \in (A \cap B' \cap C') \cup (A' \cap B \cap C') \cup (A' \cap B' \cap C)$
6. $x \in (A' \cap B') \cup (B' \cap C') \cup (A' \cap C')$
7. $x \in (A \cap B) \cup (A \cap C) \cup (B \cap C)$
8. $x \in (A \cap B \cap C') \cup (A \cap B' \cap C) \cup (A' \cap B \cap C)$
9. $x \in (A \cap B \cap C)'$
10. $x \in A \cap C \cap B'$
11. $x \in A \cap B \cap C$
12. $x \in A \cup B \cup C$
15. a) $1 - a$
    b) $1 - b$
    c) $a + b - c$
    d) $1 - c$
    e) $1 - a + c$
    f) $a - c$

13.9 Ejercicios (pág. 584)

2. a) $\frac{5}{14}$
   b) $\frac{45}{91}$
   c) $\frac{10}{91}$
   d) $\frac{36}{91}$
3. a) $\frac{23}{36}$
   b) $\frac{1}{6}$
   c) $\frac{1}{3}$
   d) $\frac{5}{36}$
   e) $\frac{13}{36}$
   f) $\frac{2}{3}$
4. $\frac{1}{4}$
5. $\frac{9}{47}$
6. $\frac{8}{47}$
7. $\frac{4}{47}$
8. a) $A/(A + B)$
   b) $B/(A + B)$
   c) $(C + 1)/(C + D + 1)$
   d) $C/(C + D + 1)$
9. a) $\frac{1}{9}$
   b) $\frac{4}{9}$
   c) $\frac{5}{9}$
   d) $\frac{8}{9}$
10. $P_0 = 1 - P(A) - P(B) + P(A \cap B)$, $P_1 = P(A) + P(B) - 2P(A \cap B)$, $P_2 = P(A \cap$
12. a) 5 a 9
    b) 45 a 46
    c) 10 a 81
    d) 36 a 55

13.11 Ejercicios (pág. 591)

1. $\{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}$
2. 1326
3. 54
4. $\{H, T\} \times \{H, T\} \times \{1, 2, 3, 4, 5, 6\}$; 24 resultados
5. $52!/(13!)^4$
6. 36
   a) 18
   b) 12
   c) 24
7. a) $13 \cdot 12 \cdot 11 \cdot 72 = 123552$ (no incluyendo tríos o cuádruples)
   b) 5148
   c) 36 (no incluyendo $10JQKA$)
   d) 4
8. a) $4\binom{13}{5}\Big/\binom{52}{5}$ b) $36\Big/\binom{52}{5}$ c) $4\Big/\binom{52}{5}$
9. a) $\dfrac{2 \cdot 98!}{(49!)^2}$ b) $\dfrac{98!}{48! \cdot 50!}$
10. $\binom{98}{48}\Big/\binom{100}{50}$
11. 16
12. n^k

13.14 Ejercicios (pág. 597)

2. a) $P(A) = \frac{3}{10}$; $P(B \mid A) = \frac{6}{11}$; $P(A \cap B) = \frac{9}{55}$

4. $$\frac{\binom{98}{48}}{\binom{100}{50} - \binom{98}{50}}$$

5. $\frac{4}{11}$

6. a) $$1 - \frac{26! \cdot 34!}{21! \cdot 39!} = 1 - \frac{\binom{34}{13}}{\binom{39}{13}} = 1 - \frac{\binom{62}{5}}{\binom{39}{5}}$$

b) $$1 - \frac{\binom{34}{13} + 5\binom{34}{12}}{\binom{39}{13}}$$ c) $$\frac{3\binom{26}{5}}{\binom{39}{5}\binom{39}{13}}$$

9. $\frac{1}{2}$

15. a) $P(A) = P(B) = P(C) = \frac{1}{2}$; $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$; $P(A \cap B \cap C) = 0$

13.18 Ejercicios (pág. 608)

1. a) $P(H, H) = p_1p_2$; $P(H, T) = p_1(1 - p_2)$; $P(T, H) = (1 - p_1)p_2$; $P(T, T) = (1 - p_1)(1 - p_2)$
   b) Si
   c) No
   d) H_1 y H_2, H_1 y T_2, H_2 y T_1, T_1 y T_2
2. a) $\frac{319}{512}$
   b) $\frac{11}{1024}$
   c) 6

3. $$\binom{10}{3}\frac{5^7}{6^{10}} = \frac{390625}{2519424}$$

4. a) $\frac{5}{16}$
   b) $\frac{1}{2}$
   c) $\frac{3}{16}$
5. a) $(5!)^2/10! = \frac{1}{252}$
   b) $\frac{1}{2}$
6. a) $36p^{10} - 80p^9 + 45p^8$
   b) $\frac{7}{128}$
7. Es ventajoso apostar cantidades iguales de dinero

8. $$\binom{n}{k}\frac{w^kb^{n-k}}{(w + b)^n}$$

9. $$\binom{8}{3}\frac{17^5}{18^8} = \frac{9938999}{1377495072}$$

10. $\frac{193}{512}$
11. $1 - (19/20)^{10} = 0{,}4013$
12. $\frac{193}{512}$
14. $59 \leq n \leq 65$

15. a) $f(p) = (1 - p)^2 + p^3$
    b) $(\sqrt{31} - 4)/3$

13.20 Ejercicios (pág. 614)

1. a) $f(k) = 2k$
   b) $f(k) = 3^k$
   c) $f(k) = p_k$, donde p_k es el k-ésimo primo ≥ 2
   d) una tal función es $f(k) = (g(k), h(k))$, donde

$$g(k) = \frac{m^2(k) + 3m(k)}{2} - k + 2, \quad h(k) = k - \frac{m^2(k) + m(k)}{2},$$

y

$$m(k) = \left[\frac{\sqrt{8k - 7} - 1}{2}\right]$$

donde $[x]$ representa el mayor entero $\leq x$
   e) $f(k) = 2^{g(k)}3^{h(k)}$, donde $g(k)$ y $h(k)$ son las definidas en (d)

13.22 Ejercicios (pág. 617)

1. $n = 0$: máx $= 1$, mín $= \frac{1}{2}$
   $n = 1$: máx $= \frac{1}{4}$, mín $= 0$
   $n = 2$: máx $= \frac{1}{8}$, mín $= 0$
   $n = 3$: máx $= \frac{1}{16}$, mín $= 0$
3. a) $1 - qp^3 - pq^3$
   b) $\frac{7}{8}$
   c) $\frac{3}{5}$
4. a) $3pq/(pq + 2)$
   b) $\frac{1}{3}$
   c) $2 \log 2 - 1$

13.23 Ejercicios variados sobre probabilidades (pág. 618)

1. $\frac{1}{12}$
2. a) $\frac{2}{19}$
   b) $\frac{1}{10}$
3. a) $\frac{1}{4}$
   b) $\frac{7}{12}$
4. a) $\frac{15}{34}$
   b) $\frac{13}{51}$
   c) $\frac{13}{165}$
5. $\frac{64}{73}$
7. 0.65
8. a) $\frac{5}{9}$
   b) $\frac{2}{9}$
   c) $\frac{2}{9}$
   d) No
9. $p^3 + 6p\left(\frac{1 - p}{2}\right)^2$
10. $np(1 - p)^{n-1} + np^{n-1}(1 - p)$
11. $\frac{n}{2^{n-1}}\left(1 - \frac{n}{2^{n-1}}\right)^{m-1}$

Capítulo 14

14.4 Ejercicios (pág. 625)

1. b) $X \leq b$

2. a) $\{\omega \mid X(\omega) \in (a, b], Y(\omega) \in (c, d]\}$
   c) $X \leq a, \quad Y \leq d$
   d) $P(a < X \leq b, c < Y \leq d) = P(X \leq b, Y \leq d) - P(X \leq a, Y \leq d) - P(X \leq b, Y \leq c) + P(X \leq a, Y \leq c)$
3. a) $\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$, $\{(5, 6), (6, 5)\}$, $\{(1, 6), (2, 5), (3, 4), (4, 3), 5, 2), (6, 1), (5, 6), (6, 5)\}$
   b) $P(X = 7) = \frac{1}{6}$; $P(X = 11) = \frac{1}{18}$; $P(X = 7$ o $X = 11) = \frac{2}{9}$
4. $Y = X_1 + X_2 + X_3 + X_4$; $P(Y = 0) = \frac{1}{16}$; $P(Y = 1) = \frac{1}{4}$; $P(Y \leq 1) = \frac{5}{16}$
5. $Y = 7X$ si $0 \leq X \leq 100$; $Y = 10X - 300$ si $X > 100$
6. a) $Z = Y - 1$
   b) $U = Y_1 + Y_2 - 1$

14.8 Ejercicios (pág. 637)

2.

t	2	3	4	5	6	7	8	9	10	11	12
$p_x(t)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

3. b) $p(-2) = \frac{1}{2}$, $p(0) = p(2) = \frac{1}{4}$
   c) $0, \frac{3}{4}, \frac{3}{4}, \frac{1}{4}, 1, 0, \frac{1}{4}, \frac{1}{4}$
5. a) $c = \frac{1}{3}$
   b) $\frac{1}{9}, \frac{1}{3}, \frac{2}{3}, \frac{8}{9}$
   c) $F(t) = 0$ para $t < 0$, $F(t) = \frac{1}{9}$ para $0 \leq t < 1$, $F(t) = \frac{1}{3}$ para $1 \leq t < 2$, $F(t) = 1$ para $t \geq 2$
   d) No hay tal t
   e) $t = 2$
6. a)

k	0	1	2	3	4
$p(k)$	$\frac{16}{81}$	$\frac{32}{81}$	$\frac{8}{27}$	$\frac{8}{81}$	$\frac{1}{81}$

$p(t) = 0$ para $t \neq 0, 1, 2, 3, 4$

   b) $F(t) = 0$ para $t < 0$, $F(t) = \frac{16}{81}$ en $[0, 1)$, $F(t) = \frac{16}{27}$ en $[1, 2)$, $F(t) = \frac{8}{9}$ en $[2, 3)$, $F(t) = \frac{80}{81}$ en $[3, 4)$, $F(t) = 1$ para $t \geq 4$
   c) $\frac{8}{27}, \frac{56}{81}$
7. b) $P(X = 0) = (1 - p)^2$; $P(X = 1) = 2p(1 - p)$
8. $p_X(k) = \dfrac{2k}{n(n + 1)}$; $F_X(t) = \dfrac{[t]([t] + 1)}{n(n + 1)}$ para $0 \leq t \leq n$, donde $[t]$ es el mayor entero $\leq t$; $F_X(t) = 0$ para $t < 0$; $F_X(t) = 1$ para $t > n$
9. $p(k) = e^{-c} \dfrac{c^k}{k!}$; $k = 0, 1, 2, 3, \ldots$; $c \geq 0$

   $p(t) = 0$ para $t \neq 0, 1, 2, 3, \ldots$
10. a) $p_X(t) = \frac{1}{2}$ en $t = -1$ y $t = +1$; $p_X(t) = 0$ para los demás valores
    $F_X(t) = 0$ para $t < -1$; $F_X(t) = \frac{1}{2}$ para $-1 \leq t < 1$; $F_X(t) = 1$ para $t \geq 1$

11. $P(A) = \frac{2}{3}$; $P(B) = \frac{2}{3}$; $P(A \cap B) = \frac{1}{3}$; $P(B \mid A) = \frac{1}{2}$; $P(A \cup B) = 1$

14.12 Ejercicios (pág. 647)

1. a) $c = 1$; $f(t) = 1$ si $0 \leq t \leq 1$; $f(t) = 0$ para los demás valores de t.
   b) $0, \frac{1}{3}, \frac{1}{3}$
2. $c = \frac{1}{2}$; $F(t) = 0$ si $t < -\pi/2$; $F(t) = \frac{1}{2}\cos t$ si $-\pi/2 \leq t < 0$; $F(t) = 1 - \frac{1}{2}\cos t$ si $0 \leq t < \pi/2$; $F(t) = 1$ si $t \geq \pi/2$
3. $c = \frac{3}{8}$; $F(t) = 0$ si $t < 0$; $F(t) = \frac{1}{4}(3t^2 - t^3)$ si $0 \leq t \leq 2$; $F(t) = 1$ si $t > 2$
4. a) $\frac{1}{2}$
   b) $\frac{1}{2}$
   c) $\frac{1}{4}$
5. a) $f(t) = 0$ si $t < \frac{1}{4}$; $f(t) = 8t - 1$ si $\frac{1}{4} \leq t < \frac{1}{2}$; $f(t) = 7 - 8t$ si $\frac{1}{2} \leq t < \frac{3}{4}$; $f(t) = 0$ si $t \geq \frac{3}{4}$
   b) $F(t) = 0$ si $t < \frac{1}{4}$; $F(t) = 4t^2 - t$ si $\frac{1}{4} \leq t < \frac{1}{2}$; $F(t) = -4t^2 + 7t - 2$ si $\frac{1}{2} \leq t < \frac{3}{4}$; $F(t) = 1$ si $t \geq \frac{3}{4}$
   c) $1, 1, \frac{1}{2}, 0, \frac{5}{16}$
6. a) $0, \frac{5}{6}, \frac{2}{3}, \frac{1}{2}$
   b) $t = 1$
7. a)

k	5	10	15	20	25	30
$P(X \geq k)$	$\frac{5}{6}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	0

   b) Cada tren *Styx* llega 5 minutos antes que un tren Lethe
8. $F_Y(t) = 0$ si $t < b$; $F_Y(t) = (t - b)/a$ si $b \leq t \leq b + a$; $F_Y(t) = 1$ si $t > b + a$
9. a) $\frac{2}{37}$
   b) $\frac{11}{37}$
   c) $\frac{36}{37}$
10. $F(t) = \dfrac{1}{2} + \dfrac{1}{\pi}\arctan\dfrac{t - b}{a}$

14.16 Ejercicios (pág. 657)

1. a) 105
   b) 10.05
2. a) $\frac{1}{2}(\sqrt{2} - 1)$
   b) $\frac{1}{4}(2 - \sqrt{2})$
   c) $\frac{1}{2}(2 - \sqrt{2})$
   d) $\frac{1}{14}(4 - \sqrt{2})$
3. a) $1 - e^{-1}$
   b) e^{-2}
   c) $(e - 1)/e^2$
   d) e^{-3}
4. $F(t) = 0$ si $t < c$; $F(t) = 1 - e^{-\lambda(t-c)}$ si $t \geq c$

6. $\lambda' = \lambda/a$, $c' = b + ac$
9. a) 0,5000
   b) 0,1359
   c) 0,9974
   d) 0,0456
10. a) 0,675
    b) 0,025
11. a) $0{,}675\sigma$
    b) $0{,}025\sigma$

12. a) 0,8185
    b) 0,8400
13. 75,98 pulgadas
14. media $= b$, varianza $= a^2$
15. $F_Y(t) = 0$ si $t < 0$; $f_Y(t) = 0$ si $t \le 0$; $f_Y(t) = e^{-t/2}/\sqrt{2\pi t}$ si $t > 0$

14.18 Ejercicios (pág. 660)

1. a) $F_Y(t) = 0$ si $t < 1$; $F_Y(t) = (t - 1)/3$ si $1 \le t \le 4$; $F_Y(t) = 1$ si $t > 4$; $f_Y(t) = \frac{1}{3}$ si $1 \le t \le 4$; $f_Y(t) = 0$ para los demás valores de t
   b) $F_Y(t) = 0$ si $t < -2$; $F_Y(t) = (t + 2)/3$ si $-2 \le t \le 1$; $F_Y(t) = 1$ si $t > 1$; $f_Y(t) = \frac{1}{3}$ si $-2 \le t \le 1$; $f_Y(t) = 0$ para los demás valores de t
   c) $F_Y(t) = 0$ si $t < 0$; $F_Y(t) = t^{1/2}$ si $0 \le t \le 1$; $F_Y(t) = 1$ si $t > 1$; $f_Y(t) = (2t)^{-1/2}$ si $0 \le t \le 1$; $f_Y(t) = 0$ para los demás valores de t
   d) $F_Y(t) = e^t$ si $t \le 0$; $F_Y(t) = 1$ si $t > 0$; $f_Y(t) = e^t$ si $t \le 0$; $f_Y(t) = 0$ si $t > 0$
   e) $F_Y(t) = e^{t/2}$ si $t \le 0$; $F_Y(t) = 1$ si $t > 0$; $f_Y(t) = \frac{1}{2}e^{t/2}$ si $t \le 0$; $f_Y(t) = 0$ si $t > 0$
   f) $F_Y(t) = 0$ si $t < 1$; $F_Y(t) = \log t$ si $1 \le t \le e$; $F_Y(t) = 1$ si $t > e$; $f_Y(t) = 1/t$ si $1 \le t \le e$; $f_Y(t) = 0$ para los demás valores de t
2. Sea ψ la función inversa de φ definida en intervalo abierto (a, b)
   Entonces $F_Y(t) = 0$ si $t \le a$; $F_Y(t) = F_X[\psi(t)]$ si $a < t < b$; $F_Y(t) = 1$ si $t \ge b$; $f_Y(t) = f_X[\psi(t)]\psi'(t)$ si $a < t < b$; $f_Y(t) = 0$ para los demás valores de t
3. a) $f_Y(t) = 0$ si $t \le 0$; $f_Y(t) = (2\pi t)^{-1/2}e^{-t/2}$ si $t > 0$
   b) $f_Y(t) = 0$ si $t < 0$; $f_Y(t) = 4t(2\pi)^{-1/2}e^{-t^4/2}$ si $t \ge 0$
   c) $f_Y(t) = 0$ si $t \le 0$; $f_Y(t) = (2\pi t^2)^{-1/2}e^{-(\log t)^2/2}$ si $t > 0$
   d) $f_Y(t) = (2\pi)^{-1/2} \sec^2 t\, e^{-(\tan^2 t)/2}$ si $|t| < \pi/2$; $f_Y(t) = 0$ si $|t| \ge \pi/2$

14.22 Ejercicios (pág. 666)

2. a) $P(X = x_1) = P(X = x_2) = P(Y = y_1) = P(Y = y_2) = \frac{1}{2}(p + q)$
   b) $p = q = \frac{1}{2}$
3. a) $F(x, y) = \left(\frac{x - a}{b - a}\right)\left(\frac{y - c}{d - c}\right)$ si $a \le x \le b$ y $c \le y \le d$,

$F(x, y) = \frac{x - a}{b - a}$ si $a \le x \le b$ e $y > d$, $F(x, y) = \frac{y - c}{d - c}$ si $x > b$ y $c \le y \le d$,

$F(x, y) = 1$ si $x > b$ e $y > d$, $F(x, y) = 0$ para los demás valores de x e y
b) $F_X(x) = (x - a)/(b - a)$ si $a \leq x \leq b$; $F_X(x) = 0$ si $x < a$; $F_X(x) = 1$ si $x < b$;
$F_Y(y) = (y - c)/(d - c)$ si $c \leq y \leq d$; $F_Y(y) = 0$ si $y < c$; $F_Y(y) = 1$ si $y < d$
c) X e Y son independientes

5. $P(Y = 0) = \frac{1}{2}$; $P(Y = 1) = P(Y = 2) = \frac{1}{4}$
7. $\frac{1}{5}, \frac{4}{5}$
8. $\frac{1}{3}$
9. b) $f(x, y) = \frac{1}{2}$ si $(x, y) \in Q$; $f(x, y) = 0$ si $(x, y) \notin Q$;
$f_X(x) = 1 - |x|$ si $|x| \leq 1$; $f_X(x) = 0$ si $|x| > 1$;
$f_Y(y) = 1 - |y|$ si $|y| \leq 1$; $f_Y(y) = 0$ si $|y| > 1$.
X e Y no son independientes
10. $g(u, v) = f(u + a, v + b)$

14.24 Ejercicios (pág. 673)

1. b) $f_V(v) = 1 + v$ si $-1 \leq v < 0$; $f_V(v) = 1 - v$ si $0 \leq v \leq 1$; $f_V(v) = 0$ si $|v| > 1$
c) U y V no son independientes
2. b) $f_V(t) = 2 - 2t$ si $0 \leq t \leq 1$; $f_V(t) = 0$ para los demás valores de t
c) U y V son independientes
3. b) $g(u, v) = ue^{-u}$ si $u > 0, 0 < v < 1$; $g(u, v) = 0$ para los demás valores de u y v
c) $f_U(u) = ue^{-u}$ si $u > 0$; $f_U(u) = 0$ si $u \leq 0$
d) $f_V(v) = 1$ si $0 < v < 1$; $f_V(v) = 0$ para los demás valores de v
4. b) $g(u, v) = (v/2\pi)e^{-(1+u^2)v^2/2}$ si $v \geq 0$
c) $f_U(u) = [\pi(1 + u^2)]^{-1}$
5. $f_Z(t) = \pi^{-1/2}e^{-t^2}$
6. b) $f_U(u) = 0$ si $u < 0$; $f_U(u) = (u/\sigma^2) \exp\left(-\frac{u^2}{2\sigma^2}\right)$ si $u > 0$; $F_U(t) = 0$ si $t < 0$;
$F_U(t) = 1 - \exp\left(-\frac{t^2}{2\sigma^2}\right)$ si $t \geq 0$

14.27 Ejercicios (pág. 681)

1. $E(X) = \frac{7}{2}$, $\text{Var}(X) = \frac{35}{12}$
7. a) $E(X) = \text{Var}(X) = \lambda$
b) No
c) $E(X) = 1/\lambda$, $\text{Var}(X) = 1/\lambda^2$
d) $E(X) = m$, $\text{Var}(X) = \sigma^2$
8. a) $C(r) = (r - 1)/2$
b) $F_X(t) = \frac{1}{2}|t|^{1-r}$ si $t < -1$; $F_X(t) = \frac{1}{2}$ si $-1 \leq t \leq 1$; $F_X(t) = 1 - \frac{1}{2}t^{1-r}$ si $t > 1$
c) $P(X < 5) = 1 - 5^{r-1}/2$; $P(5 < X < 10) = (5^{1-r} - 10^{1-r})/2$
d) X tiene esperanza finita cuando $r > 2$; $E(X) = 0$
e) Varianza finita para $r > 3$; $\text{Var}(X) = (r - 1)/(r - 3)$
9. $E(X) = E(Y) = -\frac{1}{37}$; $E(Z) = -1767/50653$; $E(X + Y + Z) = -4505/50653$
10. $E(X) \to \infty$ cuando $n \to \infty$
12. a) $(2/\pi)^{1/2}$
b) $e^{1/2}$

c) $e^2 - e$
d) $(\pi/2)^{1/2}$

14.31 Ejercicios (pág. 691)

4. 251
5. 0
6. La desigualdad de Chebyshev da $1/9$; las tablas dan 0,0027
8. b) 0,6826
9. b) 0,0796
10. a) 0,0090
 b) 0,0179

Capítulo 15

15.5 Ejercicios (pág. 703)

2. a) No
 b) Sí
 c) Sí
 d) No
 e) Sí
 f) No
 g) Sí

3. a) Ni lo uno ni lo otro
 b) Seminorma
 c) Seminorma
 d) Ni lo uno ni lo otro
 e) Seminorma
 f) Norma
 g) Ni lo uno ni lo otro
 h) Ni lo uno ni lo otro

6. a), b), c)
8. b) El polinomio de a) más $\frac{7}{8}(5x^3 - 3x)\int_{-1}^{1}(5t^3 - 3t)f(t)\,dt + \frac{9}{128}(35x^4 - 30x^2 + 3) \times \int_{-1}^{1}(35t^4 - 30t^2 + 3)f(t)\,dt$
9. $a = -3e/4 + 33/(4e)$, $b = 3/e$, $c = \frac{15}{4}(e - 7/e)$
10. $\frac{15}{128}(1 + 14x^2 - 7x^4)$
11. b) $\|f - P\|^2 = \frac{1}{96}$
12. a) $\|P - f\|^2 = \dfrac{n-1}{n} - \dfrac{\log^2 n}{n-1}$

 b) $P(x) = \left(\dfrac{12}{(n-1)^2} - \dfrac{6(n+1)}{(n-1)^3}\log n\right)x + \dfrac{4(n^3-1)\log n}{(n-1)^4} - \dfrac{6(n+1)}{(n-1)^2}$;

 $\|P - f\|^2 = 36\log 2 - 28\log^2 2 - \frac{23}{2} = 0{,}0007$ cuando $n = 2$

13 a) $\|P - f\|^2 = \dfrac{1}{2n}\left[(n-2)e^{2n} + 4e^n - n - 2\right]$

 b) $P(x) = (18 - 6e)x + 4e - 10$; $\|P - f\|^2 = 20e - \frac{7}{2}e^2 - \frac{57}{2} = 0{,}0038$

14. b) $\varphi_{k+1}(x) = \dfrac{\sqrt{(2k+1)(2k+3)}}{k+1}\,x\varphi_k(x) - \dfrac{k}{k+1}\sqrt{\dfrac{2k+3}{2k-1}}\,\varphi_{k-1}(x)$, donde $\varphi_k = P_k/\|P_k\|$

15. b) $\sum_{k=0}^{m} P_k^2(x) = \frac{p_m}{p_{m+1}} [P_m(x)P'_{m+1}(x) - P_{m+1}(x)P'_m(x)]$

15.9 Ejercicios (pág. 713)

1. a) $P(x) = \frac{1}{6}(x^2 + 13x + 12)$
   b) $P(x) = \frac{1}{2}(x^2 - 5x + 6)$
   c) $P(x) = -\frac{1}{6}(x^3 - 6x^2 + 5x - 6)$
   d) $P(x) = 2x^3 + x^2 - x - 2$
   e) $P(x) = -5x^3 - x^2 + 10x - 5$
2. $P(x) = \frac{1}{640}(9x^4 - 196x^2 + 640)$
4. a) $Q(x) = 2x^3 + 3x^2 - x - 3$
   b) $Q(x) = 4x^3 + 7x^2 - 3x - 7$
5. a) $P(32) = \frac{31}{21}$; $f(32) - P(32) = \frac{43}{42}$
   b) $P(32) = \frac{992}{255}$; $f(32) - P(32) = -\frac{709}{510}$
   c) $P(32) = \frac{43}{15}$; $f(32) - P(32) = -\frac{11}{30}$
   d) $P(32) = -\frac{403}{1530}$; $f(32) - P(32) = \frac{2114}{765}$
7. a) $L_0(x) = \frac{1}{72}(u - 1)(u - 3)(u - 4)(u - 6)$; $L_1(x) = -\frac{1}{30}u(u - 3)(u - 4)(u - 6)$;
   $L_2(x) = \frac{1}{18}u(u - 1)(u - 4)(u - 6)$; $L_3(x) = -\frac{1}{24}u(u - 1)(u - 3)(u - 6)$;
   $L_4(x) = \frac{1}{180}u(u - 1)(u - 3)(u - 4)$
   b) $P(2.6) = 20$
8. b) $x \geq 1{,}581$
   c) $h \leq 0{,}0006$
12. a) $a = 0$, $b = 1$
   b) $c = 1$, $d = -2L'_k(x_k)$
13. d) Sean $B_0(x) = 1$ y $B_n(x) = (x - x_0)(x - x_0 - nh)^{n-1}/n!$ para $n \geq 1$; el único polinomio P de grado $\leq n$ que satisface las condiciones $P(x_0) = c_0$, $P'(x_1) = c_1$, $P''(x_2) = c_2, \ldots, P^{(n)}(x_n) = c_n$ viene dado por $P(x) = c_0B_0(x) + \cdots + c_nB_n(x)$
16. x

15.13 Ejercicios (pág. 721)

4. b)

8	−5040	13068	−13132	6769	−1960	322	−28	1		
9	40320	−109584	118124	−67284	22449	−4536	546	−36	1	
10	−3628800	1026576	−1172700	723680	−269325	63273	−9450	870	−45	1

c) $1 + 2x + 2x^2 - 3x^3 + x^4$

5. c)

8	1	127	966	1701	1050	266	28	1		
9	1	255	3025	7770	6951	2646	462	36	1	
10	1	511	9330	34105	42525	22827	5880	750	45	1

d) $-1 + 6x^{(1)} + 16x^{(2)} + 9x^{(3)} + x^{(4)}$

7. a) $\frac{4}{3}n^3 + \frac{11}{2}n^2 + \frac{61}{6}n$
b) $\frac{1}{4}n^4 + \frac{5}{6}n^3 + \frac{3}{4}n^2 + \frac{1}{6}n$
c) $\frac{1}{4}n^4 + \frac{3}{2}n^3 + \frac{11}{4}n^2 + \frac{3}{2}n$
d) $\frac{1}{5}n^5 + \frac{1}{2}n^4 + \frac{1}{3}n^3 - \frac{1}{30}n$

15.18 Ejercicios (pág. 730)

2. b) $T_n(1) = n^2, \quad T_n'(-1) = (-1)^{n-1}n^2$

5. $\operatorname{sen}\theta \operatorname{sen} n\theta = \dfrac{1 - x^2}{n} T_n'(x)$; grado $= n + 1$

7. $Q(x) = x^{n+1} - 2^{-n}T_{n+1}(x)$
8. $Q(x) = -\frac{5}{2}x^4 + \frac{35}{16}x^3 - \frac{25}{32}x^2 + \frac{25}{256}x - \frac{1}{512}$
14. b) $U_0(x) = 1, \quad U_1(x) = 2x, \quad U_2(x) = 4x^2 - 1, \quad U_3(x) = 8x^3 - 4x,$
$U_4(x) = 16x^4 - 12x^2 + 1, \quad U_5(x) = 32x^5 - 32x^3 + 6x$

15.21 Ejercicios (pág. 742)

1. a) $0{,}693773 - \epsilon$, donde $0{,}000208 \leq \epsilon \leq 0{,}001667$. De esto resultan las desigualdades $0{,}6921 < \log 2 < 0{,}6936$
b) $n = 578$
2. a) $c = \sqrt{3}/3$
b) $c_1 = \dfrac{a+b}{2} + \dfrac{b-a}{2}\dfrac{\sqrt{3}}{3}, \quad c_2 = \dfrac{a+b}{2} - \dfrac{b-a}{2}\dfrac{\sqrt{3}}{3}$
3. a) $c = \sqrt{2}/2$
b) $c_1 = \dfrac{a+b}{2} + \dfrac{b-a}{2}\dfrac{\sqrt{2}}{2}, \quad c_2 = \dfrac{a+b}{2} - \dfrac{b-a}{2}\dfrac{\sqrt{2}}{2}$
4. $a = 2 + \sqrt{2}, \quad b = 2 - \sqrt{2}$
5. $c = \sqrt{\frac{3}{2}}$

7. d) $$\int_a^b f(x)\,dx = \frac{b-a}{18}\left[5f\left(\frac{b+a}{2} - \frac{b-a}{2}\sqrt{\frac{3}{5}}\right) + 8f\left(\frac{b+a}{2}\right) + 5f\left(\frac{b+a}{2} + \frac{b-a}{2}\sqrt{\frac{3}{5}}\right)\right]$$

10. a) $\log 2 = 0{,}693254 - \epsilon$, donde $0{,}000016 \leq \epsilon \leq 0{,}000521$; esto nos da las desigualdades $0{,}69273 < \log 2 < 0{,}69324$
b) $\log 2 = 0{,}69315023 - \epsilon$, donde $0{,}00000041 \leq \epsilon \leq 0{,}00001334$; esto nos conduce a las desigualdades $0{,}693136 < \log 2 < 0{,}693149$
11. d) $\log 2 = 0{,}693750 - \epsilon$, donde $0{,}000115 \leq \epsilon \leq 0{,}003704$; esto nos conduce a las desigualdades $0{,}69004 < \log 2 < 0{,}69364$

ÍNDICE ALFABÉTICO

Ch

D

E

F

G

H

I

J

K

L

M

R

S